ABSTRACT

Title of dissertation: TIMING THE STATE OF LIGHT WITH ANOMALOUS DISPERSION AND A GRADIENT ECHO MEMORY

Jeremy B. Clark, Doctor of Philosophy, 2013

Dissertation directed by: Dr. Paul D. Lett
National Institute
of Standards and Technology

Quantum information science, the great nexus of two outstanding scientific achievements of the 20^{th} century (quantum mechanics and information theory), has enjoyed accelerated experimental progress over the past two decades. More recently, researchers have begun to see quantum continuous variables as a powerful alternative to discrete variable quantum information protocols involving the use of qubits. Optical four–wave mixing in ^{85}Rb vapor has proven to be an attractive source of continuous–variable entanglement since it is capable of producing distillable multi–spatial–mode entangled EPR states of light. This self–assembled mode structure can be exploited as independent quantum channels for entanglement–based algorithms in quantum communication and quantum computation. In this thesis, we will study how this flexible quantum resource can be manipulated using atomic ensembles acting as a dispersive medium or as a storage medium.

We study the effects of anomalous dispersion on these entangled states by sending one half of the state through a gain–assisted fast–light medium and measuring

the state's quantum mutual information. We observe an advance in the maximum of the quantum mutual information between modes. In contrast, due to uncorrelated noise added by a small phase–insensitive gain, we do not observe any statistically significant advance in the "leading edge" of the mutual information. We also study the storage and retrieval of multiplexed optical signals in a Gradient Echo Memory (GEM) at relevant four–wave mixing frequencies in ^{85}Rb. Temporal multiplexing capabilities are demonstrated by storing multiple classical images in the memory simultaneously and observing the expected first–in last–out order of recall without obvious cross–talk. We also develop a technique wherein selected portions of an image written into the memory can be spatially targeted for readout and erasure on demand. The effects of diffusion on the quality of the recalled images is characterized. Our results indicate that Raman–based atomic memories may serve as a flexible platform for the storage and retrieval of multiplexed optical signals.

TIMING THE STATE OF LIGHT WITH ANOMALOUS DISPERSION AND A GRADIENT ECHO MEMORY

by

Jeremy B. Clark

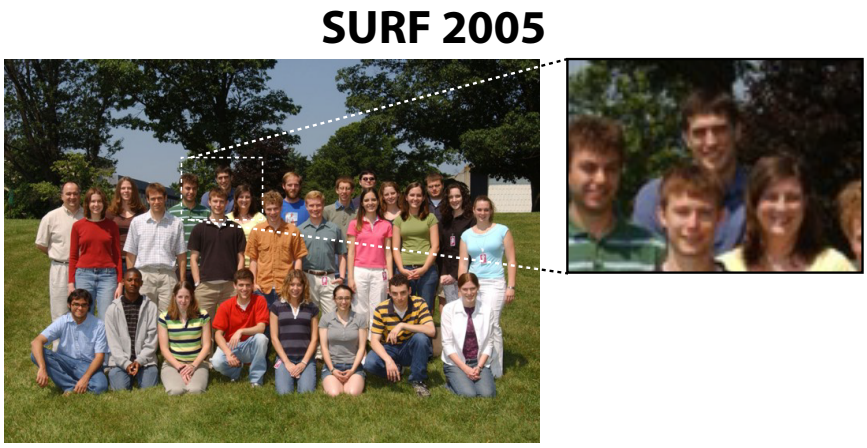
Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Dr. Paul D. Lett, Co-Chair/Off-Campus Advisor
Dr. Steven L. Rolston, Co-Chair/Campus Advisor
Dr. William D. Phillips
Dr. Luis A. Orozco
Dr. Thomas E. Murphy

To Uncle Bob.

Acknowledgments

It is probably not possible to adequately convey the depth of gratitude I feel toward those who contributed to my education and training during my years at NIST. It goes without saying that I should begin by thanking my advisor, Paul Lett, for his mentorship and friendship over the eight and a half years I've known him (see below). I think Paul struck the right balance between providing much–needed guidance and lab direction while also leaving some room for creativity on the part of his group members. His personal and scientific advice were invaluable to me over the course of my education, and I am deeply grateful to have been his student.



Summer Undergraduate Research Fellowship (SURF) program participants during the summer of 2005. This was the year that I met Paul, who taught me that lasers could be used to do more interesting things than just point at stuff.

I also feel very fortunate to have interacted with several graduate students and postdocs who profoundly enriched my experience at NIST. Starting graduate school with Neil Corzo was a uniquely rewarding experience. I learned a lot from him as we puzzled our way through quantum optics together during the early days, and I thought it was great how this habit never seemed to taper through time (although the questions seemed to get harder!). Similarly, I would like to thank Zhifan Zhou not only for his heroic efforts in the lab (particularly during late night data taking marathons), but for his steady optimism. Zhifan always kept the place cheerful, and it was certainly a noticeable loss for more than just the nonlinear optics group when he returned to China.

I regret that my graduate career had to wait until its last year to work closely with Ulrich Vogl, who I think has an extremely far–reaching physical intuition. I hope I get the chance to collaborate with Uli sometime in the future, and to learn as much from him as I did during his last year at NIST.

Fortunately for me, I was afforded a bit more time to work together with Ryan Glasser, who is without a doubt one of the smartest and most generous people I have

ever met. I also suspect I won't be having as many interesting and valuable physics conversations that involve steady use of the word "derp" without him around. I really hope that Ryan and I stay lifelong friends.

I greatly benefited from two very productive years working with Quentin Glorieux (who also taught me how to climb large, fake rocks). Quentin's style made me feel like my opinions and intuition mattered as we went along, and he was always very generous with sharing credit for our progress in the lab. I'm looking forward to following his ever–evolving interests now that he has made academic research his career (which I know will be a bright one).

I should also thank Alberto Marino, Andrew Lance, and Lincoln Turner, who all helped me learn the nuts and bolts of experimental physics very early during my time as a graduate student. In particular, Alberto's mentorship during later years (after I had gotten the hang of things a bit) was an extremely important influence on my scientific development. I regret that I was unable to have him as a member of my defense committee, which I feel would have properly reflected the magnitude of his contribution to my education.

I feel very fortunate to have met Bill Phillips, Gretchen Campbell, Kris Helmerson, Trey Porto, Ian Spielman, and all of the members of the laser cooling and trapping group (including Kim and Gail, of course!). Bill in particular is a uniquely thoughtful group leader who shows inexhaustible interest in not only in physics but the well–being of the graduate students and postdocs who help keep the wheels spinning. I should say that I feel a particularly large debt of gratitude towards Trey Porto for the considerable time and effort he spent in helping me strengthen my intuition about a whole range of topics in physics through impromptu conversations in his office. This would occasionally cost him a substantial chunk of his time, but I always felt like I left these conversations with a much clearer physical intuition.

I'd also like to express my sincere gratitude to Kevin Jones for his bottomless willingness to indulge whatever off–the–wall physics musings (or any other topic for that matter) I came to his office to discuss. Kevin surely has one of the best senses of humor out of anyone who existed anywhere ever. I always looked forward to summers at NIST when he'd be around, and I will definitely miss working with him.

Mom, Dad, and Trev: it is probably not possible to ask for a more supportive and loving family. I wholly attribute my healthy sense of skepticism and my high prioritization of education and personal growth to the way I was raised. I feel I owe every measure of any of my personal success to you.

Anna, you have doubtlessly been my greatest source of support and strength through some of the most difficult times of my graduate career and my personal life. You have shown more patience and trust in me than I ever would have thought possible. More importantly, you have taught me what it means to truly love another person, which I now understand is life's greatest joy.

Finally, I would like to thank my uncle Bob for his sustained interest in my progression in science over the course of my life. It is for this reason that this thesis is dedicated to you, Uncle Bob. I hope it makes you proud.

This work was supported by the National Science Foundation through the Physics Frontier Center of the Joint Quantum Institute.

Contents

Chapter 1: Nonclassical Light, Continuous Variables, and Entanglement

1.1 Introduction

The goal of this chapter is to introduce (in as self-contained a manner as possible) a variety of ideas that will be of central conceptual importance to this thesis. In particular, we will encounter the notion of nonclassical light fields, quantum continuous variables, squeezed electromagnetic fields, continuous–variable entanglement, and the covariance matrix (which is useful for describing Gaussian states).

1.2 Examples of "Classical States"

Consider the coherent state [1], often referred to as the "most" classical state (for reasons that will be elucidated shortly):

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{1.1}$$

where it is understood that $|n\rangle$ represents a Fock state containing n photons. It is a trivial exercise to show that

$$\begin{aligned} \langle\alpha|\, n\, |\alpha\rangle &= \langle\alpha|\, aa^\dagger\, |\alpha\rangle = |\alpha|^2 \\ \langle\alpha|\, n^2\, |\alpha\rangle &= \langle\alpha|\, aa^\dagger aa^\dagger\, |\alpha\rangle = |\alpha|^4 + |\alpha|^2 \end{aligned} \tag{1.2}$$

Therefore,

$$\Delta n = \sqrt{\langle n^2\rangle - \langle n\rangle^2} = \sqrt{\bar{n}}, \tag{1.3}$$

the hallmark of a Poisson distribution. Since Poissonian statistics can be obtained by a semiclassical treatment of photodetection [2], the coherent state fits into our simple picture of "classical state."

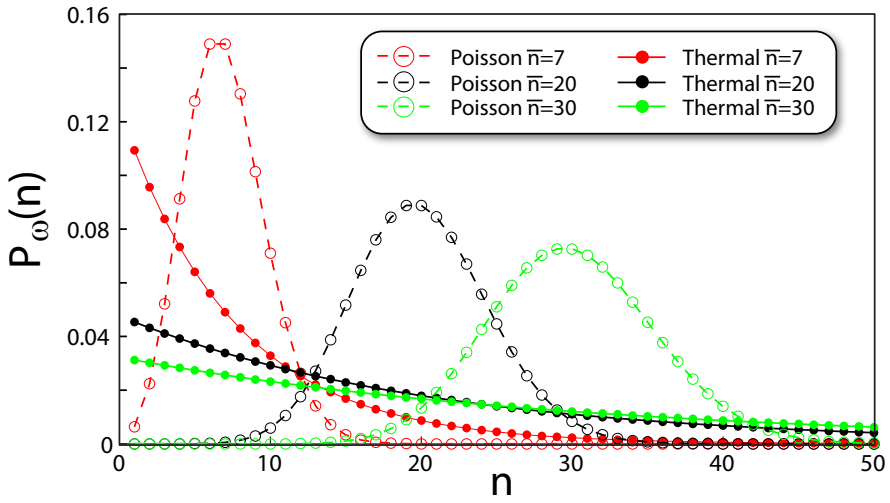


Fig. 1.1: Poisson and thermal distributions for single–mode states with average photon numbers $\langle n\rangle$ =7, 20, and 30 (red, black, and green curves, respectively).

Let us now consider a thermal state, which originates from black–body radiation [3] and whose photon populations are given by Boltzmann's Law [4]. The

photonic gas is assumed to be at thermal equilibrium:

$$P_\omega(n) = \frac{\exp(-E_n/k_BT)}{\sum_{n=0}^{\infty} \exp(-E_n/k_BT)}. \tag{1.4}$$

The energy of a given quantized level is given by

$$E_n = (n + \frac{1}{2})\hbar\omega. \tag{1.5}$$

Since we are interested in populations above vacuum, we will ignore the factor of $\frac{1}{2}$ and combine this expression with Boltzmann's law. Recalling the convergence of a geometric series, we find that

$$\begin{aligned} P_n(\omega) &= \frac{\exp(-n\hbar\omega/k_BT)}{\sum_{n=0}^{\infty} \exp(-n\hbar\omega/k_BT)} \\ &= (1 - \exp(-\hbar\omega/k_BT)) \exp(-n\hbar\omega/k_BT). \end{aligned} \tag{1.6}$$

Noting that the mean photon number is given by

$$\bar{n} = \sum_{n=0}^{\infty} nP_\omega(n) = \left(\exp(\hbar\omega/k_BT) - 1 \right)^{-1} \tag{1.7}$$

we finally arrive at the Bose–Einstein distribution describing the occupation numbers of the thermal state, $\hat{\rho}_{\text{th}}$:

$$\hat{\rho}_{\text{th}} = \sum_{n=0}^{\infty} P_\omega(n) \left|n\right\rangle \left\langle n\right| = \frac{1}{\bar{n}+1} \left(\frac{\bar{n}}{\bar{n}+1} \right)^n \left|n\right\rangle \left\langle n\right|. \tag{1.8}$$

From this expression, the photon number variance is evaluated to be

$$\begin{aligned}(\Delta n)^2 &= -\bar{n}^2 + \sum_{n=0}^{\infty} P_\omega(n) n^2 \\ &= \bar{n}^2 + \bar{n},\end{aligned} \tag{1.9}$$

indicative of a super–Poissonian distribution ($Q_M > 0$)[1]. Thermal and Poisson distributions are compared for states of several average photon numbers in Fig. 1.1.

1.3 Continuous Variables

Although photon detection reminds us of light's quantum mechanical nature, it is worth remembering that these photons are quantized excitations of a *continuously varying* electromagnetic field. As we will see in later chapters, continuous variables associated with quantum systems can also be exploited for quantum information processing[2] [6, 7]. We turn now to developing some of the most important ideas associated with quantum continuous variables (CVs), namely the notion of phase space, squeezed states, and continuous–variable entanglement.

The picture of a single mode of electromagnetic radiation prevalent before the quantum revolution was that of a traveling electromagnetic wave varying sinusoidally in time and space. For what follows, it will be convenient to use complex algebra to describe the evolution of such a plane wave $\mathbf{E}(z,t)$ polarized along $\hat{x}$ and propagating

[1] Q_M is the Mandel Q parameter. $Q_M \equiv \frac{\langle(\Delta n)^2\rangle}{\langle n\rangle} - 1$ [5].

[2] While many of these continuous–variable protocols can be qualitatively similar to their discrete variable counterparts, there are often very important differences. We will discuss these ideas at length in a later chapter.

in the z–direction:

$$\mathbf{E}(z,t) = \frac{1}{2}\left(E_0 e^{i(kz-\omega t+\phi)} + E_0^* e^{-i(kz-\omega t+\phi)}\right)\hat{x} \tag{1.10}$$

For notational convenience, we will drop the vector notation and just bear in mind that the field is polarized along $\hat{x}$. Absorbing the phase ϕ into the complex field amplitude E_0, we can re–express Eq. 1.10 using complex algebra:

$$\mathrm{E}(z,t) = |E_0|\left(\underbrace{\frac{E_0+E_0^*}{2|E_0|}}_{X}\cos(kz-\omega t) + i\underbrace{\frac{E_0-E_0^*}{2|E_0|}}_{Y}\sin(kz-\omega t)\right) \tag{1.11}$$

where we have introduced the unitless (normalized) "quadratures" X and Y. From inspection of their definitions, the quadratures denote the real and imaginary parts of the complex field amplitude, which oscillate $\pi/2$ out of phase from one another. The evolution of the field can then be described as a single point evolving around a circular trajectory in the phase space (Fig. 1.2).

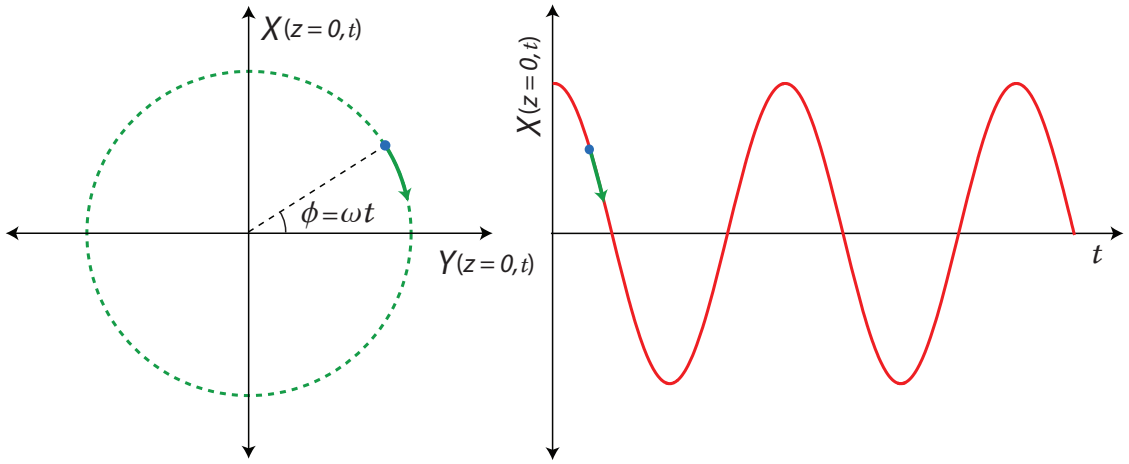


Fig. 1.2: Graphical depiction of purely wavelike monochromatic radiation in a phase space picture. The X and Y quadratures oscillate out of phase with one another at angular frequency ω.

The quantization procedure amounts to promoting these field mode amplitudes to the non-commuting creation and annihilation operators, $\hat{a}^\dagger$ and $\hat{a}$, respectively. The field amplitude associated with a single excitation of a field mode of frequency ω occupying a quantized mode volume V is given by

$$\epsilon_\omega = \sqrt{\frac{\hbar\omega}{\epsilon_0 V}}. \tag{1.12}$$

The electric field operator is given by [8]

$$\hat{\mathrm{E}}(z,t) = i\sqrt{\frac{\hbar\omega}{\epsilon_0 V}}\hat{a}e^{i(kz-\omega t)} - \mathrm{H.C.}. \tag{1.13}$$

It is convenient then to define the unitless quadrature operators

$$\hat{X} = \frac{1}{2}(\hat{a}^\dagger + \hat{a}) \tag{1.14}$$

$$\hat{Y} = \frac{i}{2}(\hat{a}^\dagger - \hat{a}). \tag{1.15}$$

such that the electric field operator at a given position ($z = 0$) can be written as

$$\hat{\mathrm{E}}(z=0,t) = \epsilon_\omega(\hat{X}\cos\omega t + \hat{Y}\sin\omega t). \tag{1.16}$$

The point of expressing the electric field operator in this way is to emphasize that the quadrature operators are *normalized* to the vacuum field strength associated with the mode of interest.

The creation and annihilation operators do not commute. By extension, the

commutator of the quadrature operators evaluates to

$$[\hat{X}, \hat{Y}] = \frac{i}{2}, \tag{1.17}$$

which carries extremely far–reaching consequences for the behavior of the quadratures. The Heisenberg uncertainty relation for two non–commuting operators can be written as [9]

$$\langle (\Delta \hat{A})^2 \rangle \langle (\Delta \hat{B})^2 \rangle \geq \left(\frac{1}{2i} \langle [\hat{A}, \hat{B}] \rangle \right)^2. \tag{1.18}$$

According to our definition of the field quadratures, the relevant uncertainty relation assumes the form

$$\langle (\Delta \hat{X})^2 \rangle \langle (\Delta \hat{Y})^2 \rangle \geq \frac{1}{16} \tag{1.19}$$

It is now clear that thinking of the electromagnetic field as a well-defined point moving through phase space (Fig. 1.2) is inadequate since both quadratures cannot be simultaneously known. The best approximation we can make to the classical phase space picture is to set both the quadrature uncertainties to the same value, with the product of the variances minimized (see Fig. 1.3a). This state is satisfied by the coherent state of Eq. 1.1. It is now clear why the coherent state is often referred to as the "most classical" state possible since it most closely mimics the behavior of a classical phase space picture.

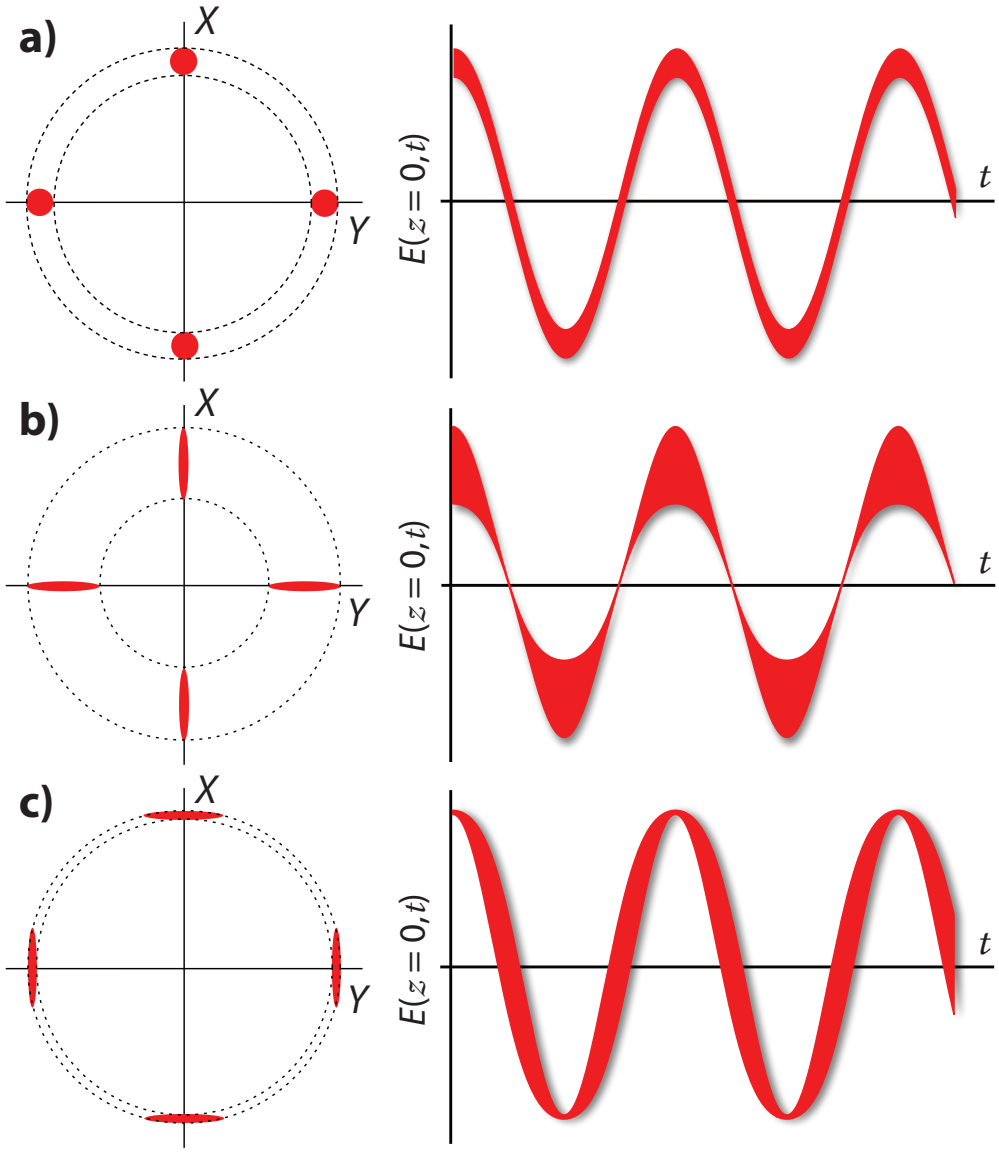


Fig. 1.3: Depiction of -5 dB squeezed states, which are routinely achievable in the laboratory. Phasor diagram interpretations are depicted on the left while the fluctuations of time–dependent electric field $E(z = 0, t)$ are depicted on the right.

1.4 Squeezed Electric Fields

Let us consider an experimentally accessible example of a nonclassical state of light. There is no reason that the fluctuations of the non–commuting quadratures need be equal (or that their product reach a minimal uncertainty). If one "squeezes" the uncertainty of one of the quadratures below the quadrature noise associated with a coherent state (the shot noise limit), the Heisenberg uncertainty principle places a lower bound on the extent to which the fluctuations of the other quadrature must increase. In essence, this behavior amounts to preserving the commutation relations of the creation and annihilation operators. Such states are sensibly referred to as "squeezed states," whose degree of squeezing is often quantified in terms of the minimum variance of the coherent state. For example, a "-5 dB squeezed state" (plotted in Fig. 1.3) is a state that exhibits a variance in its quadrature fluctuations which is $10^{-5/10} \approx 0.32$ smaller than shot noise.

If we assume a phase reference $\phi = 0$ (in the sense of Eq. 1.10), we see that squeezing the in–phase ($\hat{X}$) and out–of–phase ($\hat{Y}$) quadratures dramatically changes the character of the light field (see Fig. 1.3). When the in-phase ($\hat{X}$) quadrature is squeezed, the fluctuations of the field amplitude are suppressed. Similarly, squeezing the out–of–phase quadrature $\hat{Y}$ gives rise to suppressed fluctuations in the field's zero crossing, giving rise to a "phase squeezed" state. In general, it is necessary to define a phase reference to refer to amplitude and phase-squeezed states (and, by extension, the amplitude and phase quadratures $\hat{X}$ and $\hat{Y}$.) If we choose another reference local oscillator $\cos(\omega t - \phi)$ as our phase reference, one invokes the so–called

generalized field quadratures [3, 10]

$$\hat{X}^{\phi} = \frac{1}{2}(\hat{a}^{\dagger}e^{i\phi} + \hat{a}e^{-i\phi}) \tag{1.20}$$

$$\hat{Y}^{\phi} = \frac{i}{2}(\hat{a}^{\dagger}e^{i\phi} - \hat{a}e^{-i\phi}) \tag{1.21}$$

to express the field as

$$\hat{\mathrm{E}}(t) = \epsilon_{\omega}(\hat{X}^{\phi}\cos(\omega t - \phi) + \hat{Y}^{\phi}\sin(\omega t - \phi)). \tag{1.22}$$

1.4.1 Generation of Squeezed Light

How does one squeeze the electromagnetic field? The most general ingredient is the presence of a nonlinear optical process whose Hamiltonian is *quadratic* in $\hat{a}$ and $\hat{a}^{\dagger}$ [7]. For example, the Hamiltonian governing an optical parametric amplifier (OPA) undergoing spontaneous parametric down–conversion (which takes a very similar form to the four–wave mixing Hamiltonian we will encounter shortly) can be expressed as [11]

$$\hat{H}_{int} \propto \hat{a}^{\dagger 2}\hat{a}_{\mathrm{pump}} - \hat{a}^{2}\hat{a}^{\dagger}_{\mathrm{pump}}. \tag{1.23}$$

In the limit of a strong driving field, $\hat{a}_{\mathrm{pump}}$ is treated as a classical field[3] yielding

$$\hat{H}_{int} = i\hbar\frac{\kappa}{2}\left(\hat{a}^{\dagger 2}e^{i\theta} - \hat{a}^{2}e^{-i\theta}\right). \tag{1.24}$$

[3] Treating the pump classically amounts to treating the pump field is an undepleted coherent state. Formally, one makes the substitution [11] $\hat{a} \to |\alpha_{\mathrm{pump}}|e^{i\theta}$ where the pump phase θ is defined relative to the phase reference chosen to describe the field.

Under the undepleted pump approximation, κ is proportional to the square root of the pump photon flux (the pump intensity).

Despite this nonlinearity in the Hamiltonian, the interaction effectively yields a *linear mixing* of the operators in the Heisenberg picture. This is easy to see by examining the Heisenberg equation of motion (assuming a pump phase θ=0 for the moment):

$$\frac{d}{dt}\hat{a}(t) = \frac{-i}{\hbar}[\hat{a}(t), \hat{H}_{\text{int}}] = \kappa\hat{a}^{\dagger}(t) \tag{1.25}$$

whose solution is given by

$$\hat{a}(t) = \hat{a}(0)\cosh\kappa t + \hat{a}^{\dagger}(0)\sinh\kappa t. \tag{1.26}$$

The motion of the quadratures then trivially follows as

$$\hat{X}(t) = e^{\kappa t}\hat{X}(0) \tag{1.27}$$

$$\hat{Y}(t) = e^{-\kappa t}\hat{Y}(0). \tag{1.28}$$

Assuming an input state whose quadrature variances are equal and minimized according to the laws of quantum mechanics, the behavior of the output quadrature variances goes as

$$\langle(\Delta\hat{X}(t))^2\rangle = e^{2\kappa t} \tag{1.29}$$

$$\langle(\Delta\hat{Y}(t))^2\rangle = e^{-2\kappa t}. \tag{1.30}$$

As expected, we find that the general effect is to squeeze one quadrature at the expense of another. The linear affine transformation expressed in Eq. 1.26 is an example of a *Bogoliubov transformation* [7, 12]

$$\hat{a}_{\text{out}} = \sum_j A_{ij}\hat{a}_j + B_{ij}\hat{a}_j^\dagger + \gamma_i \tag{1.31}$$

where the matrices $\mathbf{A}$ and $\mathbf{B}$ are constrained to preserve the Bosonic commutation relations of $\hat{a}_{\text{out}}$ and $\hat{a}_{\text{out}}^\dagger$.

1.4.2 Displaced Squeezed States

It is often desirable to introduce the squeezing operator $\hat{S}(\xi)$ to derive the Bogoliubov transformation governing the field evolution. Making use of the time evolution operator

$$\hat{U} = e^{-\frac{i}{\hbar}\hat{H}(t-t_0)}, \tag{1.32}$$

the "squeeze operator" [3] follows from inserting the Hamiltonian in Eq. 1.24 into Eq. 1.32:

$$\hat{S}(\xi) = \exp\left(\frac{1}{2}(\xi^*\hat{a}^2 - \xi\hat{a}^{\dagger 2})\right). \tag{1.33}$$

The constant ξ is proportional to the relevant nonlinear susceptibility, the interaction time, and the pump strength. From the complexity of the pump's coherent amplitude α_{pump}, ξ is assumed to be complex. For notational convenience, we express the complex coefficient $\xi = re^{i\theta}$, where r is referred to as the squeezing parameter (which can assume values $0 \leq r < \infty$) and θ defines a "squeezing phase."

It is useful to define the so–called displacement operator $\hat{D}(\alpha)$ according to [1]

$$\hat{D}(\alpha) = \exp\left(\alpha\hat{a}^\dagger - \alpha^*\hat{a}\right) \tag{1.34}$$

where $\alpha = |\alpha|e^{i\phi}$ is also assumed to be complex. It is worth noting that $\hat{S}(\xi)$ and $\hat{D}(\alpha)$ do not commute, so their order of operation on any quantum state (or operator in the Heisenberg picture) is important.

In the Schrodinger picture, applying operators to the vacuum state $|0\rangle$ yields so–called displaced squeezed states denoted as

$$\hat{D}(\alpha)\hat{S}(\xi)\,|0\rangle \equiv |\alpha, \xi\rangle\,. \tag{1.35}$$

When working in the Heisenberg picture, it is helpful to invoke the Baker–Campbell–Hausdorff formula [13]

$$\begin{aligned} e^{\hat{A}+\hat{B}} &= e^{-\frac{1}{2}[\hat{A},\hat{B}]}e^{\hat{A}}e^{\hat{B}} \\ &= e^{\frac{1}{2}[\hat{A},\hat{B}]}e^{\hat{B}}e^{\hat{A}} \end{aligned} \tag{1.36}$$

to evaluate expectation values of interest. For example, one finds that

$$\hat{S}^\dagger(\xi)\hat{a}\hat{S}(\xi) = \hat{a}\cosh r - \hat{a}^\dagger e^{i\theta}\sinh r \tag{1.37}$$

$$\hat{S}^\dagger(\xi)\hat{a}^\dagger\hat{S}(\xi) = \hat{a}^\dagger\cosh r - \hat{a}e^{-i\theta}\sinh r \tag{1.38}$$

$$\hat{D}^\dagger(\alpha)\hat{a}\hat{D}(\alpha) = \hat{a} + \alpha \tag{1.39}$$

$$\hat{D}^\dagger(\alpha)\hat{a}^\dagger\hat{D}(\alpha) = \hat{a}^\dagger + \alpha^* \tag{1.40}$$

The mean photon number for the displaced squeezed state is found to be

$$\langle \hat{n} \rangle = \langle \alpha, \xi | \, \hat{a}^\dagger \hat{a} \, | \alpha, \xi \rangle \tag{1.41}$$

$$= \langle 0 | \, \hat{S}^\dagger(\xi) \hat{D}^\dagger(\alpha) \hat{a}^\dagger \hat{a} \hat{D}(\alpha) \hat{S}(\xi) \, | 0 \rangle \tag{1.42}$$

$$= \langle 0 | \, \hat{S}^\dagger(\xi) \hat{D}^\dagger(\alpha) \hat{a}^\dagger \hat{a} \hat{D}(\alpha) \hat{S}(\xi) \, | 0 \rangle \tag{1.43}$$

$$= |\alpha|^2 + \sinh^2 r. \tag{1.44}$$

Similarly, the variance is evaluated as

$$\begin{aligned} \langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2 = |\alpha|^2 \Big(e^{-2r} \cos^2(\phi - \theta/2) + e^{2r} \sin^2(\phi - \theta/2) \Big) \\ + 2 \sinh^2(r) \cosh^2(r). \end{aligned} \tag{1.45}$$

From the variance and average photon number, one may calculate Q_M (a negative

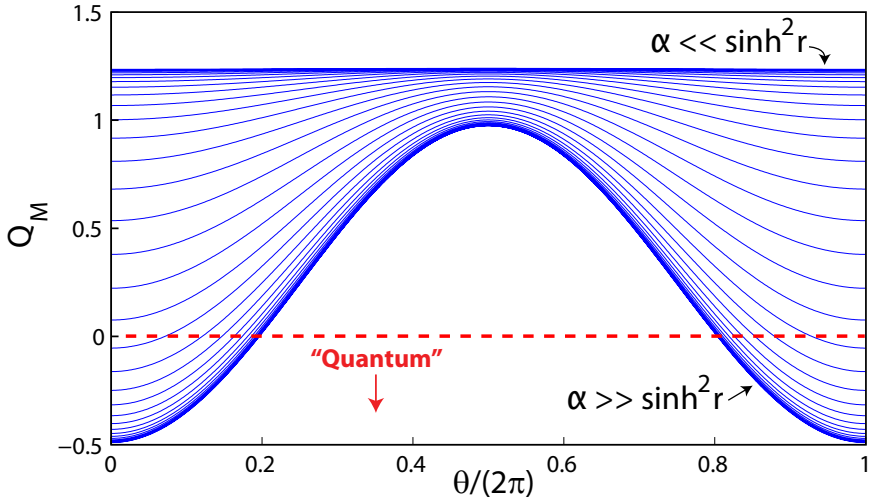


Fig. 1.4: Behavior of Q_M for a state corresponding to r=0.35 (-3 dB of squeezing) for various values of the displacement parameter α as a function the squeezing phase θ. Note that for small displacements ($\alpha \ll \sinh^2 r$), Q_M only weakly depends on the squeezing phase and never falls below 0. The state is nevertheless quantum.

value for Q_M is associated with sub–Poissonian statistics and is one signature of a

nonclassical state of light). The behavior of Q_M is non-trivial in that it depends on the squeezing parameter (r), the squeezing phase (θ), and the magnitude of the displacement (α). Figure 1.4 plots Q_M for a state with r=0.35 (which corresponds to -3 dB of continuous variable squeezing). As the plot illustrates, observing a negative Mandel Q parameter is sufficient but not necessary to conclude that a state is nonclassical. Clearly, the distinction between classical and nonclassical states is nuanced.

1.4.3 Evolution of Squeezed States

We have already computed the photon statistics of squeezed states. Let's mathematically flesh out the evolution of the electric field for a squeezed state (used to construct the plots in Fig. 1.3). Recalling that the operator describing a plane electric wave propagating along the z–axis is given by Eq. 1.22, we will explicitly carry the dependence associated with the quadrature operators at $z = 0$ according to[4]

$$\langle \hat{X}(t) \rangle = \langle \alpha, \xi | \, \hat{X}(t) \, | \alpha, \xi \rangle \tag{1.46}$$

$$\langle \sigma^2 \hat{X}(t) \rangle = \langle \alpha, \xi | \, \hat{X}(t)^2 \, | \alpha, \xi \rangle - \langle \alpha, \xi | \, \hat{X}(t) \, | \alpha, \xi \rangle^2 \tag{1.47}$$

$$\langle \hat{Y}(t) \rangle = \langle \alpha, \xi | \, \hat{Y}(t) \, | \alpha, \xi \rangle \tag{1.48}$$

$$\langle \sigma^2 \hat{Y}(t) \rangle = \langle \alpha, \xi | \, \hat{Y}(t)^2 \, | \alpha, \xi \rangle - \langle \alpha, \xi | \, \hat{Y}(t) \, | \alpha, \xi \rangle^2 \tag{1.49}$$

[4] For notational convenience, I will occasionally write $\langle \sigma^2 \hat{A} \rangle$ to denote the expectation value of the variance.

using Eqs. 1.37–1.40. The motion of these field statistics then goes as

$$\langle \hat{X}(t) \rangle = \mathrm{Re}[\alpha] \sin(\omega t) \tag{1.50}$$

$$\langle \sigma^2 \hat{X}(t) \rangle = \frac{1}{4}\Big(1 + 2\sinh^2 r - 2 \sinh r \cosh r \cos(2\omega t - \theta) \Big) \tag{1.51}$$

$$\langle \hat{Y}(t) \rangle = \mathrm{Im}[\alpha] \cos(\omega t) \tag{1.52}$$

$$\langle \sigma^2 \hat{X}(t) \rangle = \frac{1}{4}\Big(1 + 2\sinh^2 r + 2 \sinh r \cosh r \cos(2\omega t - \theta) \Big). \tag{1.53}$$

Again, r denotes the squeezing parameter and θ the squeezing angle. We also have also implicitly carried the assumption that $\alpha = |\alpha| e^{i\phi}$ is complex. Equation 1.50 conveys how the first moment of the electric field oscillates at angular frequency ω as would be expected. The field's second moment oscillates twice as fast at an angular frequency of 2ω. For $\phi = 0$ and for $\theta = 0, \pi$ these equations give rise to so–called amplitude– and phase–squeezed fields represented in Fig. 1.3. Figure 1.5 provides a more general ball-and-stick representation for a displaced squeezed state assuming arbitrary values of ϕ and θ.

So far, we have limited our consideration to squeezed states where $\alpha \neq 0$, which need not be the case. When a squeezed field does not exhibit any coherent amplitude (i.e. it is not subject to a displacement operation), one obtains the *squeezed vacuum* state which consists only of pairs of photons [3]

$$\begin{aligned} |\xi\rangle &= \hat{S}(\xi)\,|0\rangle \\ &= \frac{1}{\cosh r} \sum_{n=0}^{\infty} (-1)^n \frac{\sqrt{(2n)!}}{2^n n!} \frac{(e^{i\theta} \tanh r)^n}{\sqrt{\cosh r}}\,|2n\rangle\,. \end{aligned} \tag{1.54}$$

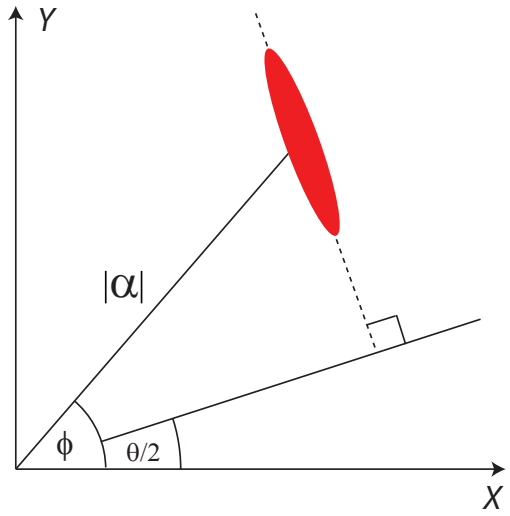


Fig. 1.5: Phase–space description of a squeezed state. The angle ϕ is given by the phase of the coherent amplitude $\alpha = |\alpha|e^{i\phi}$. The squeezing angle θ appears as a half-angle in the ball-and-stick interpretation.

As we have already seen (Fig. 1.4), the squeezed vacuum state does not exhibit sub-Poissonian statistics in its photon number given any squeezing phase. Its "nonclassicality" can be easily observed using a different kind of detection, which is sensitive to quadrature fluctuations rather than photon number fluctuations. From inspection of Eqs. 1.14 and 1.15, it is apparent that the detection of the field quadratures requires a detector which is effectively sensitive to $\hat{a}$ and $\hat{a}^\dagger$ rather than $\hat{n} = \hat{a}^\dagger\hat{a}$. To realize such a detector, it is necessary to first introduce the quantum mechanical picture of the beamsplitter.

1.5 The Beamsplitter

The beamsplitter is a linear optical element, meaning that its operator transformation equations are linear and that total photon number is preserved[5]. For the

[5] The detail about photon number preservation is important. We will shortly encounter a nonlinear interaction that (under a very convenient approximation) gives rise to an *effective* linear transformation that effectively does not preserve photon number.

following discussion, it will be helpful to refer to Fig. 1.6 to conceptually interpret the operator transformations we will be expressing. We will most often be interested in the action of a 50:50 beamsplitter, which can be described by the unitary operator

$$\hat{U} = \exp(\frac{\pi}{4}(\hat{a}_{\text{in}}^{\dagger}\hat{b}_{\text{in}} - \hat{a}_{\text{in}}\hat{b}_{\text{in}}^{\dagger})). \tag{1.55}$$

Working in the Heisenberg picture, we can express the action of this operator on the "input" creation and annihilation operators $\hat{a}_{\text{in}}$ and $\hat{b}_{\text{in}}$ as

$$\begin{pmatrix} \hat{a}_{\text{out}} \\ \\ \hat{b}_{\text{out}} \end{pmatrix} = \hat{U}^{\dagger} \begin{pmatrix} \hat{a}_{\text{in}} \\ \\ \hat{b}_{\text{in}} \end{pmatrix} \hat{U} \tag{1.56}$$

which yields (up to a common phase factor)

$$\hat{a}_{\text{out}} = \frac{1}{\sqrt{2}}(\hat{a}_{\text{in}} + \hat{b}_{\text{in}}) \tag{1.57}$$

$$\hat{b}_{\text{out}} = \frac{1}{\sqrt{2}}(\hat{b}_{\text{in}} - \hat{a}_{\text{in}}). \tag{1.58}$$

The matrix transformation equation is clearly a useful representation of the action of a beamsplitter in the Heisenberg picture. More generally, one may characterize a beamsplitter of variable reflectivity in the Heisenberg picture according to the equation [10]

$$\begin{pmatrix} \hat{a}_{\text{out}} \\ \\ \hat{b}_{\text{out}} \end{pmatrix} = \begin{pmatrix} t' & r \\ \\ r' & t \end{pmatrix} \times \begin{pmatrix} \hat{a}_{\text{in}} \\ \\ \hat{b}_{\text{in}} \end{pmatrix}. \tag{1.59}$$

To preserve the commutation relations of the output operators $\{\hat{a}_{\text{out}}, \hat{b}_{\text{out}}\}$, the fol-

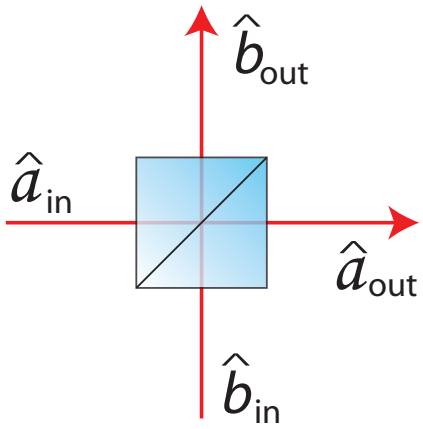


Fig. 1.6: Quantum picture of the beamsplitter. $\hat{a}_{\text{in}}$ and $\hat{b}_{\text{in}}$ represent the input annihilation operators which transform to $\hat{a}_{\text{out}}$ and $\hat{b}_{\text{out}}$.

lowing relations for the coefficients must hold

$$|r'| = |r| \tag{1.60}$$

$$|t'| = |t| \tag{1.61}$$

$$|r|^2+|t|^2 = 1 \tag{1.62}$$

$$|r'|^2+|t'|^2 = 1 \tag{1.63}$$

$$r^*t' + r't^* = 1 \tag{1.64}$$

$$r^*t + r't^* = 1. \tag{1.65}$$

Often it is convenient to treat the beamsplitter as being "phaseless," meaning that the beamsplitter does not impart any optical phase shift upon transmission. In mathematical terms, this amounts to the assumption that the transmission coefficients are purely real. Satisfying the commutation relations can then be easily achieved by constraining one reflective port of the beamsplitter to be phaseless

while assuming that the other reflective port imparts a π phase shift. Under these assumptions, we arrive at the very convenient beamsplitter transformation involving just one real adjustable parameter t:

$$\begin{pmatrix} t' & r \\ r' & t \end{pmatrix} = \begin{pmatrix} t & \sqrt{1-t^2} \\ -\sqrt{1-t^2} & t \end{pmatrix}. \tag{1.66}$$

1.5.1 *Balanced Homodyne Detection*

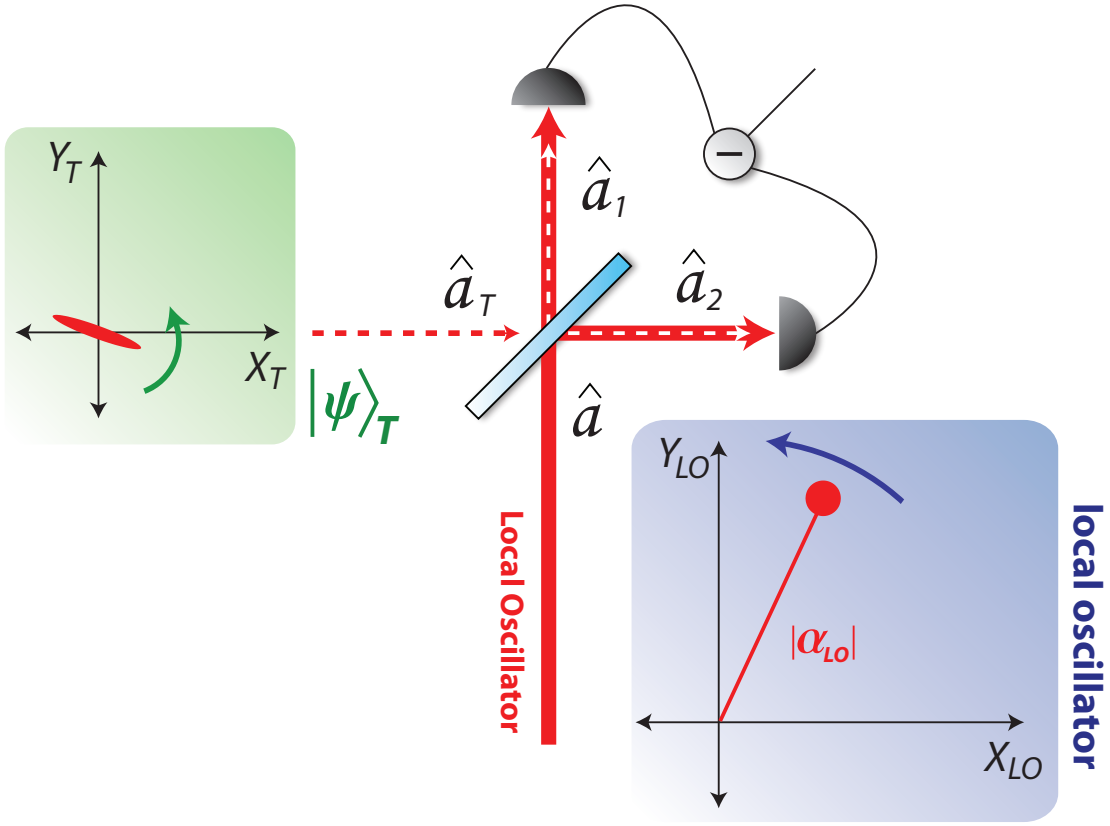


Fig. 1.7: Balanced homodyne detection. A bright local oscillator (LO) interferes with a target state $|\psi\rangle$ injected into a 50:50 beamsplitter. The rotation of the local oscillator ball–and–stick represents the angular frequency of the LO in the lab frame. The phase of the LO (ϕ) determines the detected generalized quadrature $\hat{X}^{\phi}$ of the target state.

We now return to the subject of detecting the field quadratures of a light

field. From the description of the beamsplitter in the Heisenberg picture, one can understand a very important detection technique which is sensitive to the quadrature fluctuations of the target field. Consider a coherent state $|\alpha\rangle$ (subject to creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$) and some target state of interest $|\psi\rangle_T$ (subject to creation and annihilation operators $\hat{a}_T$ and $\hat{a}_T^\dagger$) incident on a phase–free 50:50 beamsplitter (Fig. 1.7). If the spatial mode overlap of these states is perfect, they will interfere with one another to form new output modes (corresponding to Eqs. 1.57–1.58). According to Eqs. 1.57 and 1.58, subtracting the detected output modes yields the effective joint operator

$$\begin{aligned}\langle \hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2 \rangle &= \frac{1}{2}(\hat{a}_T^\dagger + \hat{a}^\dagger)(\hat{a}_T + \hat{a}) - \frac{1}{2}(\hat{a}_T^\dagger - \hat{a}^\dagger)(\hat{a}_T - \hat{a}) \\ &= \langle \hat{a}_T^\dagger \hat{a} + \hat{a}^\dagger \hat{a}_T \rangle \, .\end{aligned} \tag{1.67}$$

Recalling that $\alpha = |\alpha| e^{i\phi}$, operating on the input state $|\alpha, \psi_T\rangle$ yields

$$\begin{aligned}\langle \alpha, \psi_T | \, \hat{a}_T^\dagger \hat{a} + \hat{a}^\dagger \hat{a}_T \, | \alpha, \psi_T \rangle &= |\alpha| \langle \psi_T | \, \hat{a}_T^\dagger e^{i\phi} + \hat{a}_T e^{-i\phi} \, | \psi_T \rangle \\ &= |\alpha| \langle \hat{X}_T^\phi \rangle \, .\end{aligned} \tag{1.68}$$

Equation 1.68 illustrates the dual role of the local oscillator. For one, it provides a helpful scaling term $|\alpha|$ which effectively amplifies the signal strength. Since the field strength of realistic squeezed vacuum states is comparatively weak, this amplification factor often lifts the detected signal strength above instrumental noise floors. Perhaps more importantly, however, is the appearance of the local oscillator phase ϕ in Eq. 1.68. Modulating the relative phase Θ between the local oscillator

and the target state allows the user to detect any generalized quadrature of interest. For the state of squeezed vacuum, modulating the local oscillator phase thus permits detection of the squeezed or excessively noisy quadratures. Measuring the fluctuations of both quadratures provides access to important information about the state, which can be used to infer a variety of properties such as its purity.

1.5.2 Modeling Loss

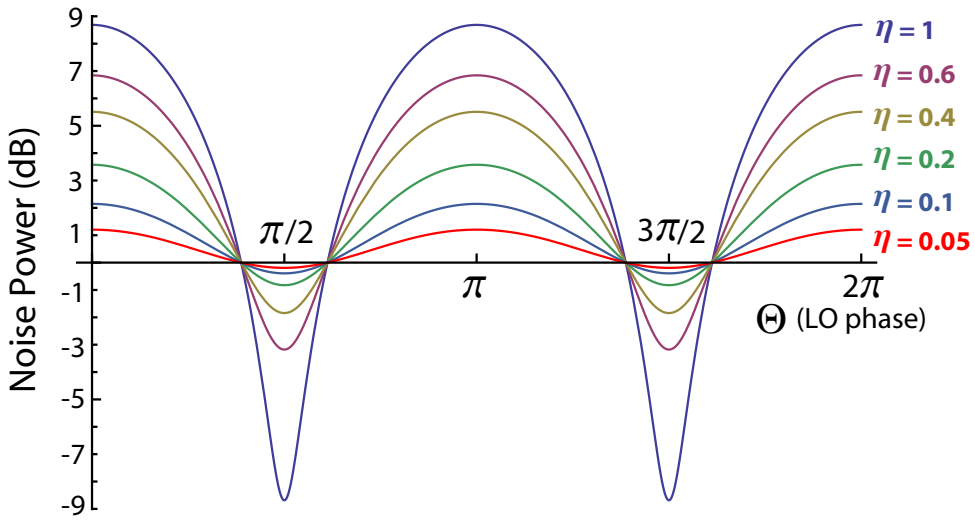


Fig. 1.8: Theoretical effect of loss on the homodyne detection of a minimum uncertainty -8.7 dB squeezed state (corresponding to a squeezing parameter $r = 1$). The plot conveys how loss yields a detection of excess noise in the unsqueezed quadrature.

Realistic experiments are subject to sources of optical loss including surface reflections and absorption, as well as effective sources of loss such as imperfect detection efficiency. These effects can often be straightforwardly modeled using the quantum description of the beamsplitter and ignoring the statistics of the detected mode (i.e. performing a partial trace). Specifically, an experiment with a net detection efficiency of η can be modeled using a beamsplitter whose transmission coefficient

$t = \sqrt{\eta}$ and whose reflection coefficient satisfies Eqs. 1.60–1.65. By convention the transmission coefficient is selected to be phaseless (i.e. purely real) since the effects due to loss (as opposed to any phase shifts) are the only effects of interest. Applying the beamsplitter Hamiltonian to the quadrature variance operator $\langle(\Delta\hat{X})^2\rangle$ of some target state of interest, the variance of the resulting quadrature operator is evaluated to be

$$\langle\sigma^2\hat{X}\rangle_{\text{out}} = \eta\,\langle(\sigma^2\hat{X}\rangle_{\text{in}} + 1 - \eta. \tag{1.69}$$

From Eq. 1.69, it is easy to see that loss linearly couples in vacuum noise while attenuating the noise properties of the target state. As illustrated in Fig. 1.8, loss can take a pure minimum uncertainty state into an impure state exhibiting excess noise in the "antisqueezed" quadrature (in comparison with the level of squeezing).

1.6 Two–Mode Squeezed States

The Hamiltonian[6] governing the evolution of two–mode squeezed states is also quadratic in creation and annihilation operators between modes $\hat{a}$ and $\hat{b}$ [11]

$$\hat{H} \propto \hat{a}^\dagger\hat{b}^\dagger e^{i\theta} - \hat{a}\hat{b}e^{-i\theta}, \tag{1.70}$$

yielding an analogous two–mode squeezing operator $\hat{S}(\xi)$:

$$\hat{S}(\xi) = e^{\xi^*\hat{a}\hat{b} - \xi\hat{a}^\dagger\hat{b}^\dagger}. \tag{1.71}$$

[6] Similar to the discussion of the OPA Hamiltonian, the phase θ is related to the pump phase in the parametric (undepleted) limit. For the case of four–wave mixing, the pump phase appears as a double angle $\theta \equiv 2\theta_{\text{pump}}$.

The complex squeezing parameter ξ is defined in an analogous way to the case of single mode squeezing, namely $\xi = re^{i\theta}$. Under the squeezing operator $\hat{S}(\xi)$, the annihilation operators of the "twin" modes undergo the transformations

$$\hat{a} \to \hat{a}\cosh r - \hat{b}^\dagger e^{i\theta}\sinh r \tag{1.72}$$

$$\hat{b} \to \hat{b}\cosh r - \hat{a}^\dagger e^{i\theta}\sinh r. \tag{1.73}$$

For an undisplaced state, it is trivial to confirm that, under these transformations,

$$\langle\hat{X}\rangle = \langle\hat{Y}\rangle = 0. \tag{1.74}$$

Although the expectation value of the first order moments of the field quadratures is zero, the modes exhibit fluctuations in excess of the shot noise limit for $r > 0$:

$$\langle(\hat{a}^\dagger + \hat{a})^2\rangle = \langle(\hat{b}^\dagger + \hat{b})^2\rangle = \cosh 2r. \tag{1.75}$$

Although such excess noise might not initially appear to contain any interesting non-classical behavior, these fluctuations are strongly correlated. To see this, it is convenient to define a set of *joint quadrature* operators $\hat{X}_-$ and $\hat{Y}_-$:

$$\hat{X}_- = \frac{1}{\sqrt{2}}(\hat{X}_a - \hat{X}_b) \tag{1.76}$$

$$\hat{Y}_- = \frac{1}{\sqrt{2}}(\hat{Y}_a - \hat{Y}_b) \tag{1.77}$$

One may easily confirm that these joint quadrature variables satisfy an analogous

commutation relation

$$[\hat{X}_-, \hat{Y}_-] = \frac{i}{2}, \tag{1.78}$$

and are therefore subject to the uncertainty relation

$$\langle \sigma^2 \hat{X}_- \rangle \langle \sigma^2 \hat{Y}_- \rangle \geq \frac{1}{16}. \tag{1.79}$$

As with the single–mode case, it possible to squeeze the fluctuations of one of these joint quadratures below the two–mode shot noise limit at the expense of the other. Using the Bogoliubov transformations of Eqs. 1.72 and 1.73, the fluctuations of the two–mode squeezed vacuum state can be straightforwardly computed as

$$\langle \sigma^2 \hat{X}_- \rangle = \cosh^2 r + \sinh^2 r + 2 \sinh r \cosh r \cos\theta \tag{1.80}$$

$$\langle \sigma^2 \hat{Y}_- \rangle = \cosh^2 r + \sinh^2 r - 2 \sinh r \cosh r \cos\theta. \tag{1.81}$$

Accordingly, selecting squeezing phase $\theta = 0$ or π alternately squeezes $\hat{Y}_-$ and $\hat{X}_-$, respectively (Fig. 1.9a).

Clearly, the two–mode squeezed vacuum state demonstrates interesting non-classical behavior since the correlations of the field quadratures are stronger than could be obtained classically. The quantum nature of the correlations between these modes is further illustrated by considering the joint state in the two–mode Fock basis $|n_a, m_b\rangle$. Beginning with the trivial relation

$$\hat{a} \, |0, 0\rangle = 0, \tag{1.82}$$

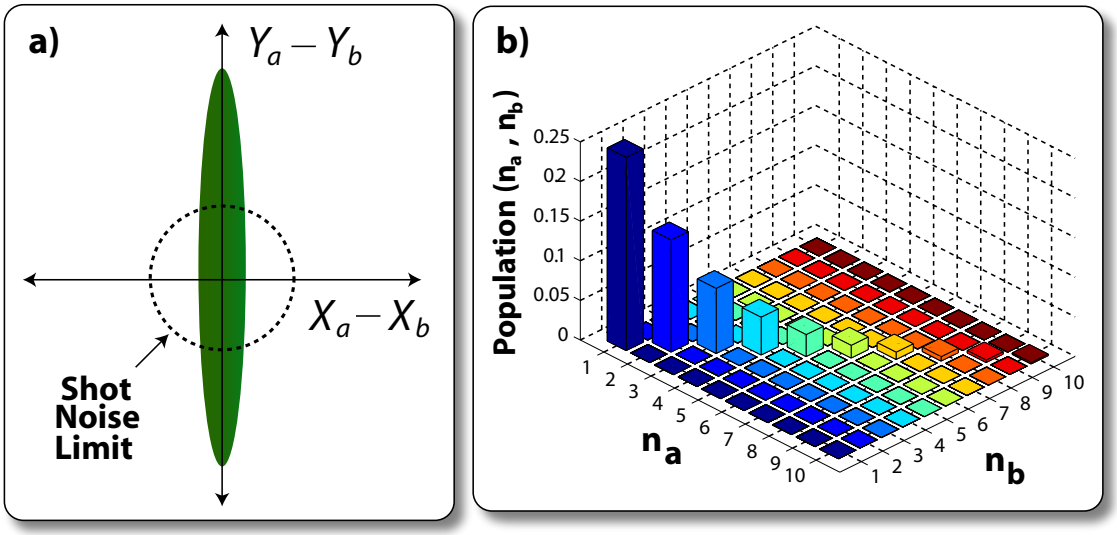


Fig. 1.9: Properties of a -8.7 dB two-mode squeezed state ($r = 1$). Panel a) illustrates the squeezing picture for a state of two–mode squeezed vacuum, and panel b) depicts the photon pair probabilities for the state.

applying the two–mode squeezing operator $\hat{S}(\xi)$ yields the equation

$$\hat{S}(\xi)\hat{a}\hat{S}^\dagger(\xi)\hat{S}(\xi)\,|0,0\rangle = (\mu\hat{a} + \nu\hat{b}^\dagger)\,|\xi\rangle \tag{1.83}$$

$$= (\mu\hat{a} + \nu\hat{b}^\dagger)\sum_{n,m} C_{nm}\,|n,m\rangle \tag{1.84}$$

$$= \sum_{n,m} C_{nm}(\mu\sqrt{n}\,|n-1,m\rangle + \nu\sqrt{m+1}\,|n,m+1\rangle) = 0. \tag{1.85}$$

Note that μ and ν are defined by Eq. 1.72. This eigenvalue equation (Eq. 1.85) is solved by the (normalized) state [3]

$$|\xi\rangle = \frac{1}{\cosh r}\sum_{n=0}^{\infty}(-1)^n e^{in\theta}\tanh^n r\,|n,n\rangle\,. \tag{1.86}$$

We have thus obtained the Schmidt decomposition [14] of the state whose Schmidt rank (number of nonzero coefficients in the decomposition) is obviously larger than

unity when $r > 0$. This indicates that the state is entangled[7]. Any ideal two–mode squeezed vacuum state exhibits perfect photon pair correlations (Fig. 1.9).

Another interesting feature worth pointing out about the two–mode squeezed state is the form of the Fock state decomposition of either of the modes upon performing a partial trace[8]. From Eq. 1.86, it is easy to see that

$$\hat{\rho}_a = \hat{\rho}_b = \sum_{n=0}^{\infty} \frac{\tanh^{2n} r}{\cosh^2 r} \left|n\right\rangle \left\langle n\right| . \tag{1.87}$$

Noting that $\langle \hat{n}_a \rangle = \langle \hat{n}_b \rangle \equiv \bar{n} = \sinh^2 r$, we can express Eq. 1.87 as

$$\hat{\rho}_{a,b} = \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(1+\bar{n})^{n+1}} \left|n\right\rangle \left\langle n\right| , \tag{1.88}$$

which is precisely the form of a thermal state (Eq. 1.8) with an effective temperature T_{eff} (defined by setting $\langle \hat{n} \rangle = \sinh^2 r$ equal to Eq. 1.7):

$$T_{\text{eff}} = \frac{\hbar\omega}{2k_B \ln(\coth r)}. \tag{1.89}$$

1.6.1 Detecting Continuous–Variable Quantum Correlations

Although it would be in principle possible to confirm the entanglement of a two–mode squeezed state by detecting photon number correlations using number–

[7] This type of entangled state is often referred to as an EPR state for historical reasons. In the limit of infinite two–mode squeezing, a two–mode squeezed state would exhibit strong correlations akin to the position/momentum correlations of the state expressed in Eq. 9 of the EPR paper [15]. An infinitely squeezed state is not physical, however, since it cannot be normalized and has an infinite energy.

[8] For any bipartite state $\hat{\rho}_{ab}$, the partial trace is evaluated by tracing over either of the states. Specifically, $\hat{\rho}_a = \text{Tr}_b[\hat{\rho}_{ab}]$ and $\hat{\rho}_b = \text{Tr}_a[\hat{\rho}_{ab}]$.

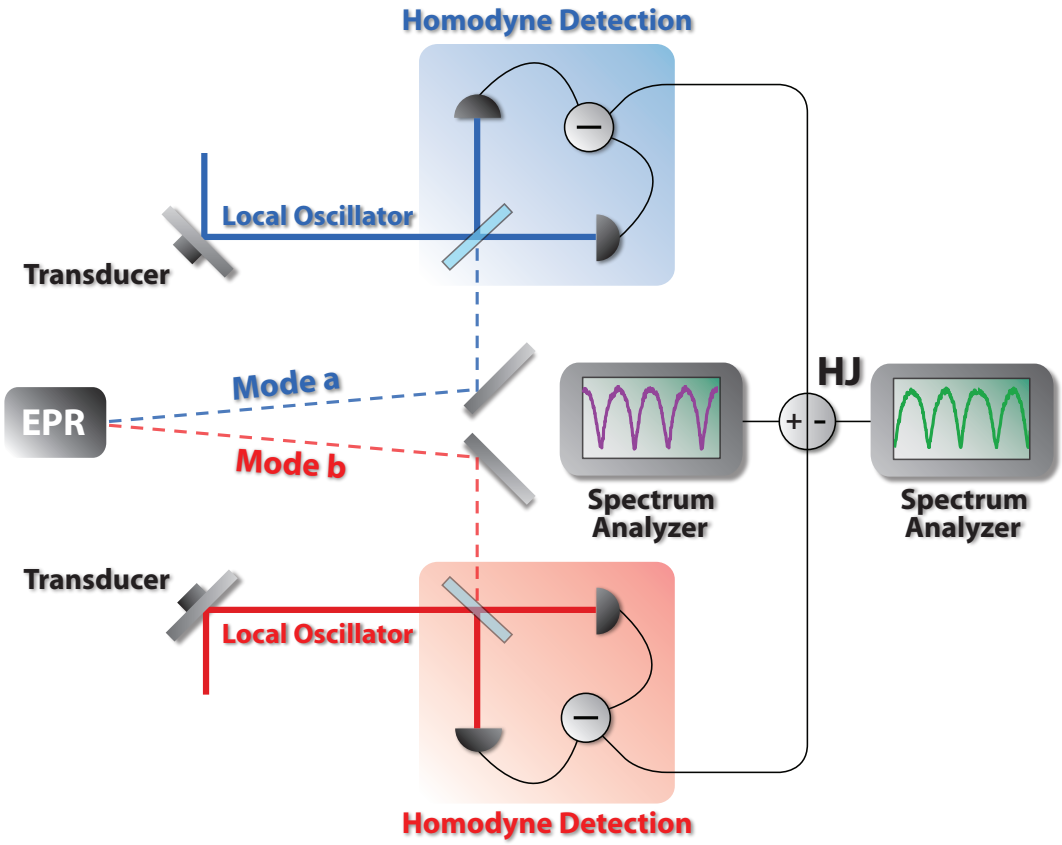


Fig. 1.10: Dual Homodyne detection scheme. Two local oscillators are prepared to match modes a and b in order to perform two simultaneous homodyne detections of the EPR fields. The output of each homodyne detector is fed into a hybrid junction (HJ) which passively provides the sum and difference signals of the homodyne detections. These sum and difference signals are recorded on a pair of spectrum analyzers as the local oscillator phases are scanned with two transducers.

resolving detectors, in this thesis we will be concerned with continuous–variable correlations of the field quadratures. Accordingly, the relevant detection technique involves performing simultaneous homodyne detections of both EPR modes (Fig. 1.10). When the local oscillator phases are selected such that quadrature correlations are being detected, subtracting the signals of these homodyne detections yields a noise power that is below the shot noise limit.

1.6.2 Inseparability Parameter, $\mathcal{I}$

Although a variety of interesting continuous–variable entanglement measures exist, we now introduce a particularly powerful entanglement witness referred to as the inseparability parameter $\mathcal{I}$, which is suitable for the states relevant to this thesis. For the two–mode squeezed states we have been discussing (even when they are subject to optical loss[9]), knowing the inseparability parameter is both necessary and sufficient to determine whether or not the state is entangled. It is convenient to define the *commuting* pair of EPR variables

$$\hat{X}_- = \frac{1}{\sqrt{2}}(\hat{X}_a - \hat{X}_b) \tag{1.90}$$

$$\hat{Y}_+ = \frac{1}{\sqrt{2}}(\hat{Y}_a + \hat{Y}_b), \tag{1.91}$$

[9] Provided that the loss is unheralded! Detecting the removal of a single photon, for example, changes the state such that the inseparability criterion is merely sufficient for entanglement.

which can be simultaneously squeezed below the shot noise limit (since they commute). The relevant inseparability parameter $\mathcal{I}$ is then defined according to

$$\mathcal{I} = \langle \sigma^2 \hat{X}_- \rangle_{\min} + \langle \sigma^2 \hat{Y}_+ \rangle_{\min} \tag{1.92}$$

where $\langle \sigma^2 \hat{X}_- \rangle$ and $\langle \sigma^2 \hat{Y}_+ \rangle$ are varied with respect to the phase of the local oscillator. Typically, the parameter is reported in a linear scale normalized to the shot noise limit. An inseparability parameter $\mathcal{I} < 2$ is necessary and sufficient to conclude that the state is entangled [16].

1.6.3 Effect of Asymmetric Loss

We have already encountered the effects of loss on a minimum uncertainty single mode squeezed state (Fig. 1.8). It can be easily shown that symmetric loss on both correlated modes of a two–mode squeezed state qualitatively leads to the same behavior as in the single–mode case (namely the persistence of squeezing and entanglement given any $\eta > 0$). Asymmetric loss between the correlated modes, however, can lead to a loss of squeezing and entanglement. Assuming an initially pure two–mode squeezed state, lossless detection of one mode and a net detection efficiency η of the other mode yields two–mode squeezing $\mathcal{R}_{ab}$ (normalized to two–mode vacuum noise) which goes as

$$\mathcal{R}_{ab} = \frac{\langle \sigma^2 \hat{X}_- \rangle}{\mathrm{SNL}} \tag{1.93}$$

$$= \frac{1}{2}\Big((1+\eta)\cosh 2r + 1 - \eta - 2\eta \sinh 2r \Big). \tag{1.94}$$

Figure 1.11 illustrates the behavior of the inseparability parameter $\mathcal{I}$ as a function of the initial two–mode squeezing and loss on one of the modes. For higher initial levels of squeezing, the level of asymmetric loss required to lose entanglement decreases.

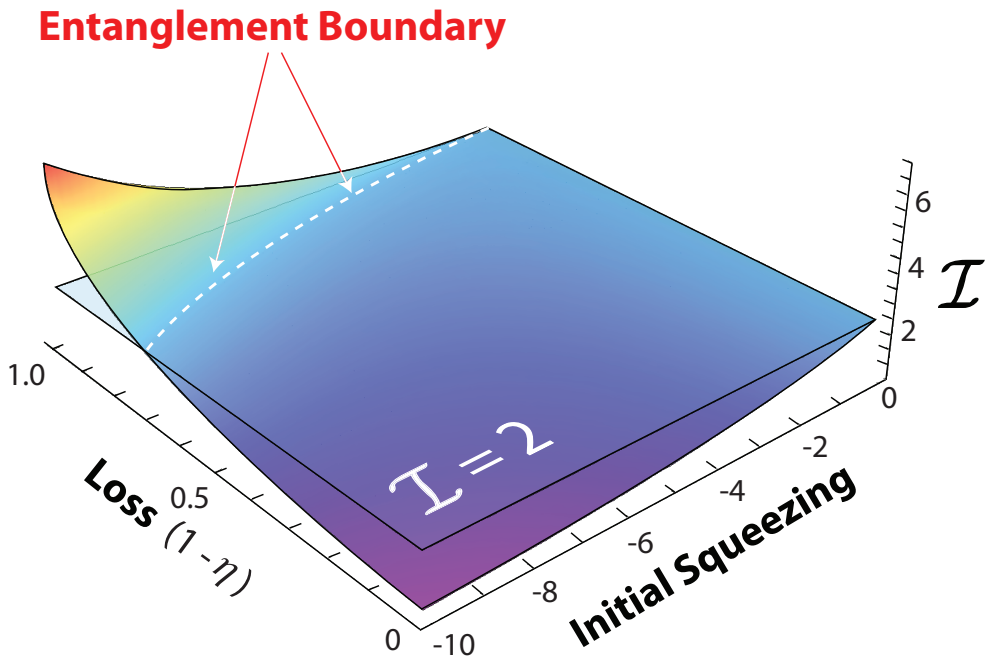


Fig. 1.11: Effect of asymmetric loss on the inseparability parameter $\mathcal{I}$. Sufficient loss on one of the correlated modes causes the state to cross into a region of state separability ($\mathcal{I} > 2$). As illustrated by the figure, this "entanglement boundary" depends on both the loss and the amount of squeezing initially present.

1.7 Four–Wave Mixing in ^{85}Rb

As previously discussed, the necessary ingredient to produce two–mode squeezed states is the existence of a nonlinear response of a dielectric medium in the presence of driving electromagnetic fields. For the experiments discussed in this thesis, the relevant technique involves a four–wave mixing (4WM) process in a vapor of ^{85}Rb

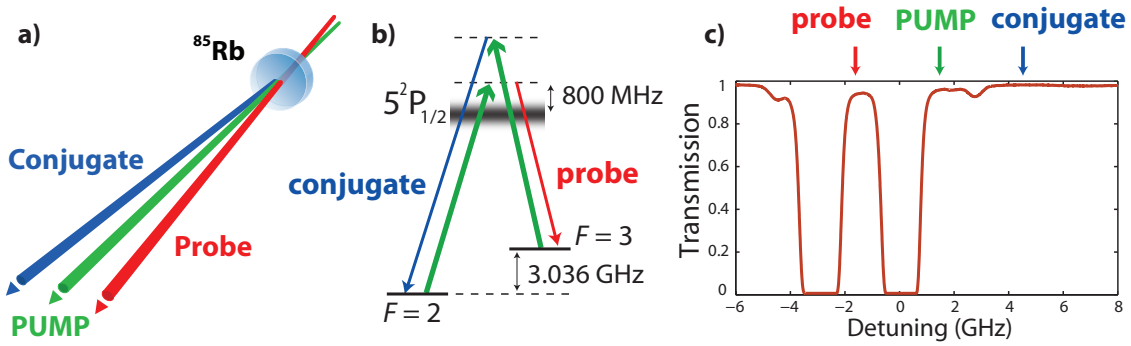


Fig. 1.12: Four-wave mixing in a vapor of ^{85}Rb. a) A strong pump beam ($\approx$ 200-400 mW) is focused to a $\approx$ 1 mm waist inside a 12 mm cell of ^{85}Rb vapor. A red-detuned probe beam is injected at an angle to the pump of approximately 0.5°. The interaction between the probe and pump leads to a gain in the probe power and stimulates the generation of a blue-detuned conjugate. b) Energy level diagram illustration of the Λ-scheme. c) Detunings of the probe, pump, and conjugate fields (indicated by the red, green, and blue arrows, respectively) relative to the measured transmission profile of the vapor (T$\approx$ 110 °C) when interrogated by a weak ($\approx$ 100 μW) beam (orange curve).

atoms heated to $\approx$100°C (Fig. 1.12)[10]. A pump beam (typically 200-500 mW) is focused to a diameter of $\approx$ 1 mm inside a vapor cell (1.2 cm in length) and detuned by approximately 800 MHz to the blue of the D_1 resonance. For some experiments, the process is "seeded" with a weak ($\approx$ 100 μW) "probe" beam focused to a diameter of $\approx$ 600 μm and detuned by 3.036 GHz to the red of the pump beam. Depending on the pump power and optical depth of the vapor, the process leads to gain ($G \approx$ 5–15 under typical conditions) in the probe power and the generation of a blue–detuned "conjugate" beam. The generated conjugate beam exhibits a detuning of 3.036 GHz to the blue of the pump beam, which is consistent with energy conservation. When only driven with the pump beam (in other words, when "vacuum–seeded"), the four–wave mixing process produces a conical emission of two–mode squeezed vacuum. The quantum correlations and continuous–variable entanglement associated

[10] The first experimental demonstration of squeezed light was also accomplished using 4WM in a vapor of Na atoms in an optical cavity [17].

with this conical emission can be detected using the dual homodyne technique we have previously described.

1.7.1 Multi–Spatial–Mode Properties

An interesting property of these twin fields is their rich spatially–dependent structure. Although this feature of the twin modes will not be developed at length in this thesis, it is worth briefly describing since it allows for interesting quantum protocols such as parallel quantum information transport [18] or holographic teleportation [19].

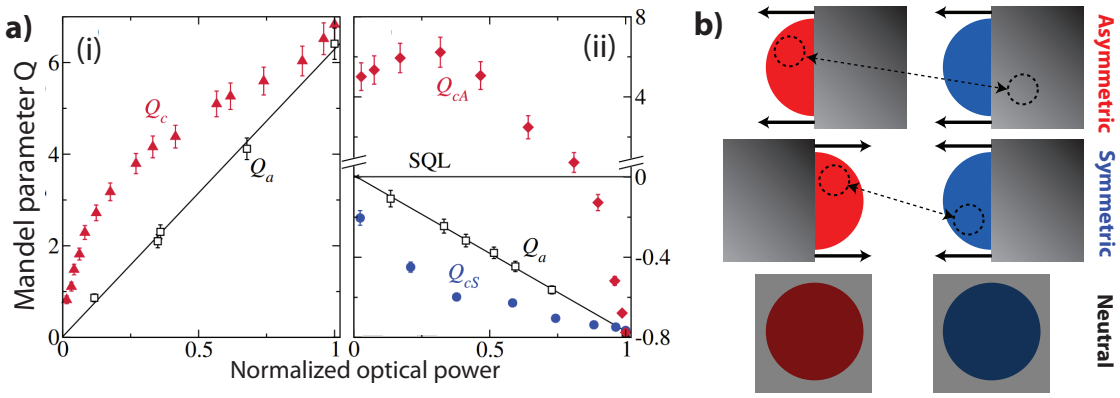


Fig. 1.13: Multi–spatial–mode properties of twin beams from four–wave mixing. Part a) includes measured plots (taken from [20]) of the Mandel Q parameter (i) for the probe as a function of the transmission for partial blocking of the probe beam. The straight line denotes the noise effects of uniform attenuation (Q_a) with a neutral density filter. The data in (ii) includes measured Mandel Q parameters Q_{cA} and Q_{cS} for asymmetric (red diamonds) and symmetric (blue dots) transverse obstruction of the detected fields, respectively. Part b) illustrates (from top to bottom) asymmetric obstruction, symmetric obstruction, and uniform attenuation of the beams with a neutral density filter. The dotted circles denote correlated coherence areas associated with the fields.

The multimode character of these beams may be observed in the non–trivial behavior of the intensity–difference noise when subject to spatially–dependent at-

tenuation (Fig. 1.13). When the twin modes are evenly attenuated using a neutral density filter, increasing the attenuation linearly decreases any intensity–difference squeezing until the shot–noise limit is reached [20]. As Fig. 1.13 illustrates, however, asymmetrically attenuating the twin beams very quickly leads to a loss of squeezing and then to the observation of excess noise. Finally, symmetric attenuation yields a loss of squeezing at a slower rate than neutral attenuation.

This behavior can be understood as resulting from the spontaneous assembly of *coherence areas* [21], pairwise–correlated subregions within the transverse profiles of the beams [20]. Each coherence area only exhibits quantum correlations with its matching coherence area positioned symmetrically opposite to the pump (Fig. 1.13b). Ultimately, these coherence areas arise due to phase–matching conditions [22], which are related to momentum conservation between the amplified probe and the stimulated conjugate. Their transverse width is related to both phase matching and the diffraction limit associated with the waist of the pump field creating the gain region inside the cell [23].

The question arises as to whether the multi–spatial–mode character of these twin beams can be exploited for imaging or metrological applications. It has been proposed [24] and demonstrated [25] that multimode squeezed vacuum fields can be used to enhance optical resolution of an object interrogated by a light field. We extended this idea by performing an experiment in which we generated a multi–spatial–mode state of vacuum–squeezed twin beams and used it to interrogate a binary intensity mask with the conjugate field. The intensity mask imprints its shape onto the conjugate field passing through it, and the remaining portion of the

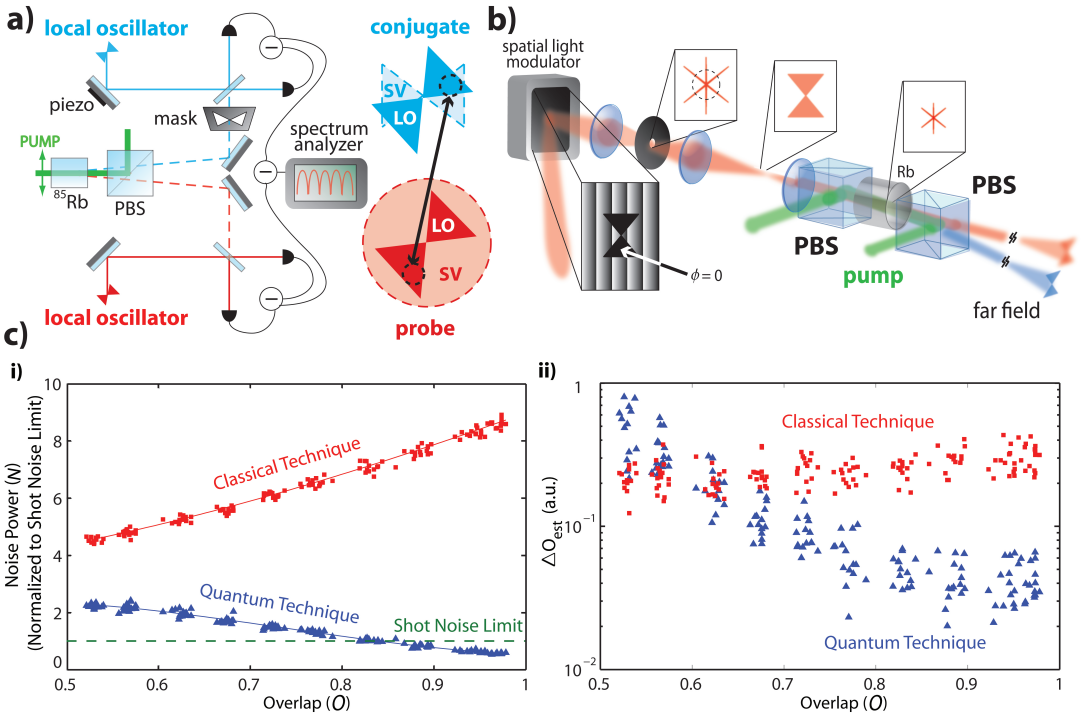


Fig. 1.14: a) Experimental set–up for imaging an object using only quantum noise. The conjugate field is directed through a binary intensity mask whose shape is to be estimated. The probe and conjugate fields are each detected with separate homodyne detectors. b) The transverse mode shape of the local oscillators is dynamically prepared using a spatial light modulator to write a spatially–dependent diffraction grating. c) Data acquired from the homodyne detections. i) When the noise of the conjugate is detected, matching the transverse profile of the local oscillator (the overlap) to the mask shape results in an increase in excess noise ("classical technique"). When both beams are detected, an increased mode matching (overlap) results in increased two–mode squeezing. ii) At high overlaps, the quantum technique is more sensitive to changes in overlap.

conjugate is sent to a homodyne detector (Fig. 1.14a).

As we have seen, the conjugate field (taken by itself) exhibits quadrature fluctuations in excess of the shot noise limit similar to a thermal state. We exploited this property of the state [26] by simulating a dynamical mode matching process between the conjugate field and the local oscillator using a spatial light modulator (SLM) to manipulate the transverse profile of the local oscillators (Fig. 1.14b). As the mode matching (overlap) improved between the local oscillator and the remaining conjugate field, the excess noise power N detected by the homodyne detection increased (Fig. 1.14c-i). We defined the sensitivity of estimating the mode overlap, O, according to the error propagation formula

$$\Delta O_{\text{est}} = \frac{\Delta N}{|\frac{\partial N}{\partial O}|}. \tag{1.95}$$

Here ΔO denotes the standard deviation of the measured overlaps (and ΔN the standard deviation of the noise power). In this way, we showed that is possible to infer the shape of the mask using only the quadrature fluctuations of the conjugate field to within some quantifiable sensitivity (Fig. 1.14c-ii). Since these fluctuations mimic those of a thermal state, we referred to this estimation technique as the "classical technique."

Although detecting the conjugate beam alone is sufficient to allow an estimation of the mask shape, it is also possible to use the correlated fluctuations of the probe beam to enhance this estimation [27]. To do so, we exploit the fact that, given appropriately chosen local oscillator phases [26], the difference signal of the

homodyne detections is squeezed. As illustrated in Fig. 1.14b, we simultaneously shaped the transverse mode profile of the local oscillators to ensure that only correlated coherence areas were selected for detection. To provide sufficient statistical sampling, we locked the relative phases of the local oscillators using active feedback based on the technique outlined in [28]. As the mode overlap between the shaped conjugate field and the local oscillators increased, we observed the two–mode noise power drop below the shot noise limit (Fig. 1.14c-i). We referred to this estimation technique as the "quantum technique."

Interestingly, we observed an increase in the sensitivity in estimating the mode overlap between the local oscillator and leaked conjugate field at high overlaps, allowing us to perform a better estimation of the mask shape. While the effects that account for this difference are subtle, a simple theoretical treatment assuming a particular mask shape and spatial mode decomposition suggests [27] that multi–spatial–mode squeezed fields would be expected to provide an advantage as we observed.

1.7.2 Phase–Insensitive Amplification

We have already seen that, ideally, a vacuum–seeded 4WM process displays perfect photon number correlations (Eq. 1.86). Let's consider the state of the twin beams when the four–wave mixing process is seeded with a coherent state $|\alpha\rangle$. We will model the behavior of the probe and conjugate fields with the single mode annihilation operators $\hat{a}$ and $\hat{b}$, respectively. We have already encountered the Bo-

goliubov transformations for the probe and conjugate operators:

$$\hat{a}_{\rm out} = \mu \hat{a}_{\rm in} + \nu \hat{b}^{\dagger}_{\rm in} \tag{1.96}$$

where for notational simplicity we denote $\cosh r = \mu$ and $\nu = -e^{i\theta} \sinh r$. It is easy to confirm that the output operators $\hat{a}_{\rm out}$ and $\hat{a}^{\dagger}_{\rm out}$ obey the necessary commutation relations:

$$\begin{aligned} [\hat{a}_{\rm out}, \hat{a}^{\dagger}_{\rm out}] &= |\mu|^2 [\hat{a}_{\rm in}, \hat{a}^{\dagger}_{\rm in}] + |\nu|^2 [\hat{b}^{\dagger}_{\rm in}, \hat{b}_{\rm in}] \\ &= |\mu|^2 - |\nu|^2 = 1. \end{aligned} \tag{1.97}$$

The transformation of Eq. 1.96 is the hallmark [29] of a *phase–insensitive amplifier* (PIA). PIAs are noisy amplifiers in the sense that they necessarily decrease the signal–to–noise ratio (SNR) of the state of light sent into the amplifier. If we define the SNR[11] according to

$$\mathrm{SNR} \equiv \frac{\langle \hat{n} \rangle^2}{\langle \sigma^2 \hat{n} \rangle} \tag{1.98}$$

we find that the SNR for a coherent state is $|\alpha|^2$. The SNR of a coherent state sent through a phase–insensitive amplifier can be computed using the relations

$$\langle \hat{a}^{\dagger}_{\rm out} \hat{a}_{\rm out} \rangle = |\mu|^2 |\alpha|^2 + |\nu|^2 \tag{1.99}$$

$$\langle \hat{b}^{\dagger}_{\rm out} \hat{b}_{\rm out} \rangle = |\nu|^2 |\alpha|^2 + |\nu|^2 \tag{1.100}$$

$$\langle \hat{a}^{\dagger}_{\rm out} \hat{a}_{\rm out} \hat{a}^{\dagger}_{\rm out} \hat{a}_{\rm out} \rangle = |\mu|^4 |\alpha|^4 + |\nu|^4 + |\mu|^2 |\nu|^2 (1 + 4|\alpha|^2). \tag{1.101}$$

[11] Often the SNR is defined as the square root of the right–hand side of Eq. 1.98, but I will observe the convention adopted by similar work [30].

From Eqs. 1.99 and 1.100, it is reasonable to associate $|\mu|^2$ with the gain G of the amplifier while $|\nu|^2$ is related to spontaneous emission. From Eq. 1.101, we find that

$$\mathrm{SNR}_{\mathrm{out}} = \frac{|\mu|^4|\alpha|^4+|\nu|^4+2|\mu|^2|\nu|^2|\alpha|^2}{|\mu|^2|\nu|^2(1+2|\alpha|^2)}. \tag{1.102}$$

From our definition of the gain G, the output SNR can be re–expressed as

$$\mathrm{SNR}_{\mathrm{out}} = \frac{G^2|\alpha|^4+(G-1)^2+2G(G-1)|\alpha|^2}{G(G-1)(1+2|\alpha|^2)}. \tag{1.103}$$

Defining the noise figure (NF) of the amplifier according to

$$\mathrm{NF} \equiv \frac{\mathrm{SNR}_{\mathrm{in}}}{\mathrm{SNR}_{\mathrm{out}}}, \tag{1.104}$$

we find that for large seed powers ($\alpha \gg 1$) and large gain ($G \gg 1$),

$$\mathrm{NF} \approx 2. \tag{1.105}$$

It is in this sense that the PIA is noisy[12].

Despite the noisy quality of the amplifier, it is a straightforward matter to compute the fluctuations in the photon number difference between the probe and conjugate given the singly seeded input state $|\alpha, 0\rangle$.

$$\langle(\Delta(\hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b}))^2\rangle = |\alpha|^2. \tag{1.106}$$

[12] A noiseless amplifier is sensitive to the input phase and permits amplification such that NF=1. While interesting, they are not terribly relevant to this thesis, so I will not discuss them.

This result is hardly surprising since the 4WM process creates or annihilates probe and conjugate photons in pairs. Therefore, the expectation value of the photon number difference should reflect the number fluctuations of the seed beam (in this case, a coherent state). If we compute the ratio R_{ab}[13] of this expected number difference to the average number difference for two coherent states (with the same average photon number), we find

$$\begin{aligned} R_{ab} &= \frac{\langle(\Delta(\hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b}))^2\rangle_{\text{4WM}}}{\langle(\Delta(\hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b}))^2\rangle_{\text{CS}}} \\ &= \frac{|\alpha|^2}{\langle\hat{a}^\dagger\hat{a}\rangle + \langle\hat{b}^\dagger\hat{b}\rangle} \\ &= \frac{|\alpha|^2}{|\alpha|^2\cosh 2r + 2\sinh^2 r} \end{aligned} \tag{1.107}$$

whose behavior is plotted in Fig. 1.15. As Fig. 1.15 illustrates, for any given squeezing parameter r, increasing α degrades the intensity–difference squeezing as should be expected.

1.8 Quasi–Probability Distributions

It is often convenient to define quadrature probability distributions (given a certain quantum state) in a manner analogous to a classical probability density. For certain quantum states, this can lead to difficulties. For instance, classical probability densities are normalized and defined to be positive everywhere, which will not necessarily be true for certain definitions of quadrature probability densities. Nevertheless, it is still possible to define a *quasi–probability* distribution for quan-

[13] Note that this is distinct from $\mathcal{R}_{ab}$ written down in Eq. 1.93.

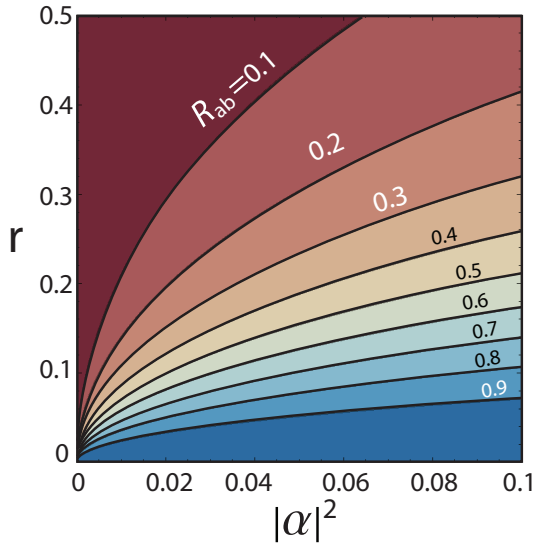


Fig. 1.15: Plot of R_{ab} versus $|\alpha|^2$ and r. The plot illustrates how, given any squeezing parameter r, increasing the seed power $|\alpha|^2$ lowers the intensity–difference squeezing.

tum states which mimics many aspects of classical probability distributions. For example, it is possible to create a distribution function that behaves much like a classical probability densities under a limited class of manipulations (e.g. computing marginal distributions or coordinate rotations). Ultimately, a quasi–probability distribution can be viewed simply as an equivalent representation of a quantum state, much like its density matrix.

1.8.1 The Wigner Function

These criteria are satisfied by the Wigner function, which is defined as[14] [31]

$$W(X,Y) = \int_{-\infty}^{\infty} \exp(iYq)\,\langle X - q/2|\,\hat{\rho}\,|X + q/2\rangle\,dq. \tag{1.108}$$

[14] A brief point regarding notation: I have tried to include hats to distinguish operators from variables, which often appear together in the relevant definitions. Probability distributions carry variables (not operators) as arguments, but evaluating the quasiprobability distributions often requires the use of operators.

For notational simplicity, I will suppress the limits for the remaining integrals in this section (with the understanding that they go from $-\infty \to \infty$). Since the density matrix behaves like an operator, it should come as little surprise that the Wigner function can be defined for operators just by replacing the density matrix with the operator of interest. This observation leads to a very important property of Wigner functions which is captured by the overlap formula

$$\mathrm{Tr}(\hat{O}_1\hat{O}_2) = \int\int W_1(X,P)W_2(X,P) \tag{1.109}$$

where $W_i(X,P)$ is the Wigner function associated with operator $\hat{O}_i$ where $i \in \{1,2\}$. This can be simply proven:

$$\begin{aligned}
\int\int W_1(X,Y)W_2(X,Y) &= \int\int\int\int e^{iY(q_1+q_2)} \langle X - q_1/2| \hat{O}_1 |X + q_1/2\rangle \\
&\quad \times \langle X - q_2/2| \hat{O}_2 |X + q_2/2\rangle \, dq_1 dq_2 dX dY \\
&= \int\int \langle X - q/2| \hat{O}_1 |X + q/2\rangle \\
&\quad \times \langle X - q/2| \hat{O}_2 |X + q/2\rangle \, dq dx \\
&= \int\int \langle q'| \hat{O}_1 |q''\rangle \langle q''| \hat{O}_2 |q'\rangle \, dq' dq'' \\
&= \int \langle q'| \hat{O}_1\hat{O}_2 |q'\rangle \, dq' \\
&= \mathrm{Tr}(\hat{O}_1\hat{O}_2)
\end{aligned} \tag{1.110}$$

So we immediately see that, upon computing the Wigner function of the state and the operator of interest, we are able to compute expectation values of observables

$$\mathrm{Tr}(\hat{\rho}\hat{O}) = \int\int W(X,Y)W_O(X,Y)dXdY. \tag{1.111}$$

It is also straightforward to compute a state's purity μ

$$\mu \equiv \mathrm{Tr}(\hat{\rho}^2) = \int\int W(X,Y)^2 dXdY, \tag{1.112}$$

or density matrix elements in any discrete basis $\sum_b |b\rangle \langle b| = \mathbb{1}$

$$\begin{aligned} \langle b'| \hat{\rho} |b\rangle &= \mathrm{Tr}(\hat{\rho} |b\rangle \langle b'|) \\ &= \int\int W(X,Y)W_{b',b}(X,Y)dXdY. \end{aligned} \tag{1.113}$$

Hopefully it is clear that the Wigner function is more than a graphically interesting way of expressing the quantum state of the electromagnetic field (Fig. 1.16). The Wigner function is a very useful representation of the state!

As it happens, there exists an uncountably infinite family of quasiprobability distributions (referred to as s–parameterized distributions).

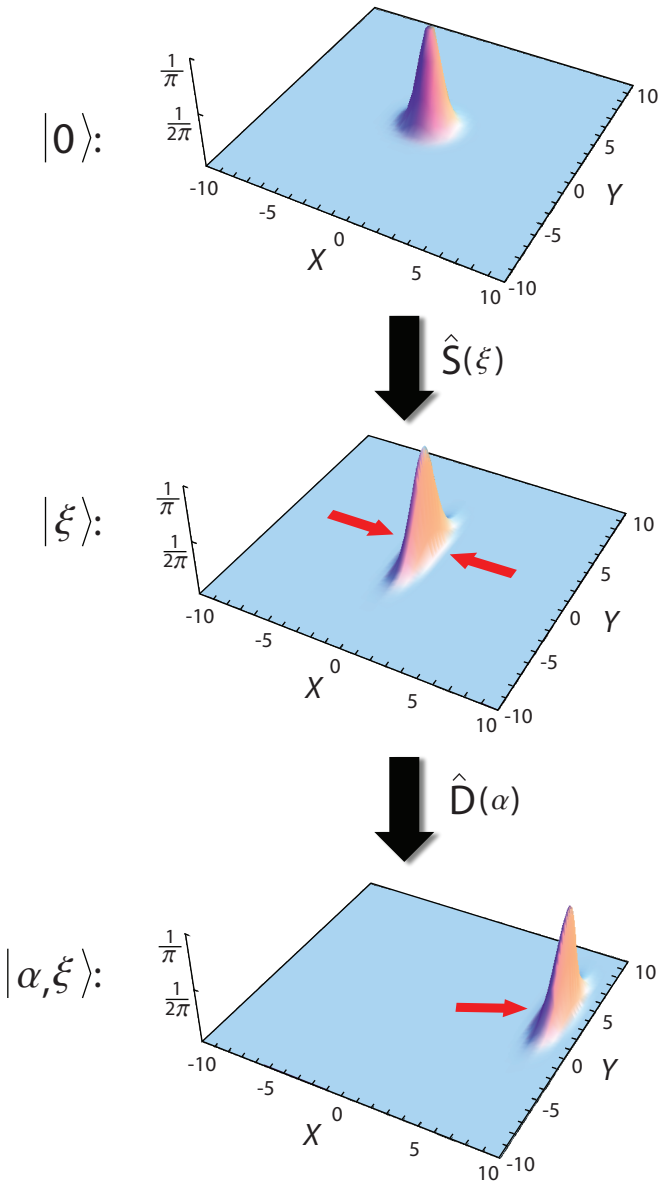


Fig. 1.16: Action of the squeezing operator $\hat{S}(\xi)$ and displacement operator $\hat{D}(\alpha)$ on the vacuum state $|0\rangle$ in the phase-space representation to yield a displaced squeezed state $|\alpha, \xi\rangle$.

1.9 Weyl Correspondence and the Covariance Matrix

To understand the physical significance of the covariance matrix, it is helpful to begin by introducing the characteristic function [32]:

$$\begin{aligned}\tilde{W}(u,v) &\equiv \int\int W(X,Y)e^{-iuX-ivY}dXdY \\ &= \mathrm{Tr}(\hat{\rho}e^{-iu\hat{X}-iv\hat{Y}})\end{aligned} \tag{1.114}$$

Note that the second step makes use of the overlap relation (Eq. 1.110).

We now note a particularly important property of states whose Wigner functions assume a Gaussian distribution. From the definition of the characteristic function (Eq. 1.114), it is straightforward to confirm the relation

$$\begin{aligned}\mathrm{Tr}[\hat{\rho}(\lambda\hat{X}+\mu\hat{Y})^k] &= i^k\frac{\partial^k}{\partial\zeta^k}\tilde{W}(\zeta\lambda,\zeta\mu)|_{\zeta=0} \\ &= \int\int W(X,Y)(\lambda X+\mu Y)^k\mathrm{d}X\mathrm{d}Y.\end{aligned} \tag{1.115}$$

Expanding these expressions and comparing powers of λ and μ yields the so–called *Weyl correspondence* principle [33–35]

$$\mathrm{Tr}[\hat{\rho}S(\hat{X}^m\hat{Y}^n)] = \int\int W(X,Y)X^mY^n\mathrm{d}X\mathrm{d}Y \tag{1.116}$$

where the symmetrization operator S has the effect shuffling any composite operator of interest into symmetric form[15].

We now apply the Weyl correspondence principle to better characterize two–

[15] For example, $S(\hat{X}\hat{Y}) = \frac{1}{2}(\hat{X}\hat{Y}+\hat{Y}\hat{X})$, etc.

mode squeezed states. For what follows, we will assume that the first moments of the twin modes that comprise the two–mode squeezed state of interest are zero[16]. Doing so helps tidy the notation but conceptually costs us very little since the main features of interest (entanglement, von Neumann entropy, etc.) do not depend on the first moments [36]. Additionally, it will be convenient to define a single column vector of operators

$$\hat{\xi} = \begin{pmatrix} \hat{X}_p \\ \hat{Y}_p \\ \hat{X}_c \\ \hat{Y}_c \end{pmatrix}$$

where $\hat{X}_{p,c}$ and $\hat{Y}_{p,c}$ are the quadrature operators for the probe and conjugate, respectively. Using the symmetrization property of the Weyl correspondence principle, we can express the quantity

$$\mathrm{Tr}[\hat{\rho}(\hat{\xi}_i\hat{\xi}_j + \hat{\xi}_j\hat{\xi}_i)/2] = \int W(\xi)\xi_i\xi_j d^{2N}\xi \equiv \gamma_{ij} \tag{1.117}$$

where we have defined the *covariance matrix*, $\boldsymbol{\gamma}$. For Gaussian Wigner functions, this relation is satisfied by [37]

$$W(\xi) = \frac{1}{(2\pi)^2\sqrt{\det\boldsymbol{\gamma}}} e^{-\frac{1}{2}\xi\boldsymbol{\gamma}^{-1}\xi^T}. \tag{1.118}$$

In other words, the covariance matrix *completely characterizes* any Gaussian state.

[16] It might be worth mentioning that this allows us to write quantities such as $\langle\sigma^2\hat{X}\rangle$ as $\langle\hat{X}^2\rangle$ since $\langle\hat{X}\rangle = 0$. When the first order moments do not vanish, it is common to express the covariance matrix in terms of variances and first moments [36].

Intuitively, this should not be surprising at all since the first and second moments are sufficient to characterize any Gaussian distribution.

We can invoke the symmetry of two–mode squeezed vacuum states to write the covariance matrix in the following "standard form" [38]

$$\boldsymbol{\gamma} = \begin{pmatrix} \langle \hat{X}_p^2 \rangle & 0 & \frac{\langle \hat{X}_p \hat{X}_c + \hat{X}_c \hat{X}_p \rangle}{2} & 0 \\ 0 & \langle \hat{Y}_p^2 \rangle & 0 & \frac{\langle \hat{Y}_p \hat{Y}_c + \hat{Y}_c \hat{Y}_p \rangle}{2} \\ \frac{\langle \hat{X}_p \hat{X}_c + \hat{X}_c \hat{X}_p \rangle}{2} & 0 & \langle \hat{X}_c^2 \rangle & 0 \\ 0 & \frac{\langle \hat{Y}_p \hat{Y}_c + \hat{Y}_c \hat{Y}_p \rangle}{2} & 0 & \langle \hat{Y}_c^2 \rangle \end{pmatrix} \tag{1.119}$$

where $\{\hat{X}_p, \hat{Y}_p\}$ and $\{\hat{X}_c, \hat{Y}_c\}$ denote the probe and conjugate quadratures, respectively. For experimentalists, it is convenient to write the covariance matrix in terms of readily measurable observables with homodyne detection. To do so, let us express the "amplitude–difference" squeezing as

$$\begin{aligned} R_{pc}^X &= \frac{1}{2} \langle (\hat{X}_p - \hat{X}_c)^2 \rangle \\ &= \frac{1}{2} \langle \hat{X}_p^2 \rangle + \frac{1}{2} \langle \hat{X}_c^2 \rangle - \frac{1}{2} \langle \hat{X}_p \hat{X}_c + \hat{X}_c \hat{X}_p \rangle \end{aligned} \tag{1.120}$$

and the "phase–sum" squeezing as

$$\begin{aligned} R_{pc}^Y &= \frac{1}{2} \langle (\hat{Y}_p + \hat{Y}_c)^2 \rangle \\ &= \frac{1}{2} \langle \hat{Y}_p^2 \rangle + \frac{1}{2} \langle \hat{Y}_c^2 \rangle + \frac{1}{2} \langle \hat{Y}_p \hat{Y}_c + \hat{Y}_c \hat{Y}_p \rangle . \end{aligned} \tag{1.121}$$

Continuing with the assumption that we are dealing with squeezed vacuum (i.e. $\langle \hat{X}_p \rangle = \langle \hat{X}_c \rangle = 0$) and writing the total excess noise (meaning noise greater than

that of the shot noise limit) of the amplitude and phase quadratures as

$$\langle \hat{X}^2 \rangle \equiv \frac{\langle \hat{X}_p^2 \rangle + \langle \hat{X}_c^2 \rangle}{2} \tag{1.122}$$

$$\langle \hat{Y}^2 \rangle \equiv \frac{\langle \hat{Y}_p^2 \rangle + \langle \hat{Y}_c^2 \rangle}{2}, \tag{1.123}$$

the covariance matrix can be expressed as

$$\boldsymbol{\gamma} = \begin{pmatrix} \langle \hat{X}_p^2 \rangle & 0 & \langle \hat{X}^2 \rangle - R_{pc}^X & 0 \\ 0 & \langle \hat{Y}_p^2 \rangle & 0 & \langle \hat{Y}^2 \rangle + R_{pc}^Y \\ \langle \hat{X}^2 \rangle - R_{pc}^X & 0 & \langle \hat{X}_c^2 \rangle & 0 \\ 0 & \langle \hat{Y}^2 \rangle + R_{pc}^Y & 0 & \langle \hat{Y}_c^2 \rangle \end{pmatrix}.$$

For a pure minimum uncertainty state, straightforward use of the Bogoliubov transformations confirms that the covariance matrix assumes the form

$$\boldsymbol{\gamma} = \begin{pmatrix} \cosh 2r & 0 & \sinh 2r & 0 \\ 0 & \cosh 2r & 0 & -\sinh 2r \\ \sinh 2r & 0 & \cosh 2r & 0 \\ 0 & -\sinh 2r & 0 & \cosh 2r \end{pmatrix} \tag{1.124}$$

where r is the usual squeezing parameter. The utility of the covariance matrix will become much clearer in later chapters when quantities such as the mutual information of the state become important.

Chapter 2: Fast Light, Classical Information, and Causality

"There was a young lady named Bright,

Whose speed was far faster than light;

She started one day,

In a relative way,

And returned on the previous night."

–Arthur Reginald Buller

2.1 Dispersion and Group Velocity

We begin by considering the propagation of electromagnetic radiation through non–conducting (dielectric) media. In this section we decompose all wave forms in the plane wave basis, which assumes that all solutions have a harmonic time dependence $e^{i\omega t}$. Often, the response of the medium is modeled as uniform, instantaneous, and isotropic by introducing $\mathbf{D} = \epsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$ into Maxwell's equations [39]:

$$\begin{aligned} \nabla \times \mathbf{E} - i\omega\mathbf{B} &= 0 \\ \nabla \times \mathbf{B} + i\omega\mu\epsilon\mathbf{E} &= 0. \end{aligned} \tag{2.1}$$

For most dielectrics (we will be confining our attention to atomic vapors), $\mu \approx \mu_0$, and we will only concern ourselves with the behavior of the electric field.

By taking the curl of Eqs. 2.1, we arrive at the Helmholtz equation

$$(\boldsymbol{\nabla}^2 + k^2)\mathbf{E} = 0, \tag{2.2}$$

where $k^2 = \mu\epsilon\omega^2$. For the remaining discussion, we will assume that any radiation is moving along the x-direction and polarized along $\hat{\boldsymbol{z}}$ so that we can simplify the notation a bit. The D'Alembert solution $u(x,t)$ for a wave oscillating at frequency ω, the most general solution to the wave equation [40], is given by

$$u(x,t) = ae^{ikx-i\omega t} + be^{-ikx-i\omega t}. \tag{2.3}$$

The velocity at which the D'Alembert solution moves, referred to as the phase velocity v_p, goes as

$$v_p = \frac{\omega}{k} = \frac{c}{\sqrt{\mu\epsilon}} = \frac{c}{n}. \tag{2.4}$$

In Eq. 2.4, c is the phase velocity of light in vacuum and n denotes the index of refraction, which modifies the wave's propagation speed[1] The important point to draw from Eq. 2.4, however, is that any variation of ϵ with frequency ω will cause different frequencies of the waveform to propagate at different velocities, leading to distortion of a propagating pulse [41].

From the basic D'Alembert solutions of Eq. 2.3, we may build general solutions

[1] Of course, it is equally valid to invert Eq. 2.4 to solve for the k-dependence of $\omega(k)$, which is generally the approach taken by the literature.

of the form

$$u(x,t) = \int A(k) e^{i(kx - \omega(k)t)} dk \tag{2.5}$$

where

$$A(k) = \frac{1}{2\pi} \int u(x,0) e^{-ikx} dx. \tag{2.6}$$

If a traveling "wave packet" being considered is sufficiently narrowband, $A(k)$ will be sharply peaked about some central value k_c, and the k–dependence of the angular frequency ω can be approximated as

$$\omega(k) \approx \omega_c + \left.\frac{\partial \omega}{\partial k}\right|_c (k - k_c) \tag{2.7}$$

by truncating its Taylor expansion. Then the traveling wave packet can be approximated as

$$\begin{aligned} u(x,t) &\approx \frac{1}{2\pi} e^{i(k_c \left.\frac{\partial \omega}{\partial k}\right|_c - \omega_c)t} \int A(k) e^{i(x - \left.\frac{\partial \omega}{\partial k}\right|_c t)k} dk \\ &= \underbrace{u(x - t\left.\tfrac{\partial \omega}{\partial k}\right|_c, 0)}_{\text{traveling wave}} \underbrace{e^{i(k_c \left.\frac{\partial \omega}{\partial k}\right|_c - \omega_c)t}}_{\text{overall phase factor}} \end{aligned} \tag{2.8}$$

where the second line of Eq. 2.8 has made use of Eq. 2.6. This is a tidy, revealing result: as long as Eq. 2.7 is a good approximation, the waveform travels largely undistorted[2] at the group velocity [39]

$$v_g = \left.\frac{\partial \omega}{\partial k}\right|_c \tag{2.9}$$

[2] Strictly speaking, any dispersion *always* causes distortion of the wave packet.

up to an overall phase factor.

It is often more convenient to express the group velocity as a function of frequency. Starting with the relation $\omega(k) = \frac{ck}{n(k)}$, we find that

$$\begin{aligned} \frac{\partial \omega}{\partial k} &= \frac{c}{n} - \frac{ck}{n^2}\frac{\partial n}{\partial k} \\ &= \frac{c}{n} - \frac{\omega}{n}\frac{\partial \omega}{\partial k}\frac{\partial n}{\partial \omega}. \end{aligned} \tag{2.10}$$

Solving for $\frac{\partial \omega}{\partial k}$ and evaluating at the central (carrier) frequency ω_0, we arrive at the frequency–dependent group velocity

$$v_g = \frac{c}{n(\omega_c) + \omega_c \frac{\partial n}{\partial \omega}\big|_c}. \tag{2.11}$$

It's worth emphasizing the fact that the group velocity is evaluated at the carrier frequency ω_0. For a given dielectric medium, the group velocity will change with carrier frequency according to Eq. 2.11 provided that the bandwidth of the incident pulse is sufficiently narrow that the approximations discussed above are valid [42].

Equation 2.11 is very suggestive: the group velocity is given by the speed of light in vacuum divided by a term that includes both the index of refraction and the derivative of the index. Accordingly, the denominator is often referred to as the group index of refraction n_g and can assume values that are larger or smaller than unity, giving rise to "slow" and "fast" light, respectively. The derivative in the denominator characterizes the medium's dispersion, the variation of a traveling

wave's phase velocity as a function of frequency[3]. It is this slope that often acts as the dominating term in determining the deviation of the group index from unity.

When $\partial_\omega n < 0$, the dispersion is said to be *anomalous* because the index of refraction typically increases with frequency over the majority of most dielectric response functions not in the neighborhood of a resonance [39]. For sufficiently negative values of $\partial_\omega n$, the group index n_g can drop below unity or even assume negative values [43]. A dielectric medium satisfying the condition

$$n(\omega_c) + \omega_c \frac{\partial n}{\partial \omega}\bigg|_c \tag{2.12}$$

is typically referred to as a "fast–light" medium.

2.1.1 The Kramers–Kronig Relations

For weakly driven systems, it is often a good approximation to model the response of a system to an external drive as being linear with a non–zero response time. This behavior can be captured by the general expression [42]

$$R(t) = \int_{-\infty}^{\infty} \mathrm{d}t' G(t - t') S(t') \tag{2.13}$$

where $S(t)$ denotes the stimulus to the system and $R(t)$ the time–dependent response. Although Eq. 2.13 seems more realistic than an instantaneous response, it

[3] To clear up a potentially confusing point: the variation of the group index with frequency is often casually referred to as "dispersion" in the literature. This term should more properly be referred to as the "group velocity dispersion" as opposed to just the "dispersion," which refers to the variation in the refractive index with frequency.

is subject to severe restrictions to be physically meaningful. Given any arbitrary response kernel $G(t-t')$, it is generally possible that the system can respond before the onset of the signal at $t = t'$ (i.e. $R(t) \neq 0$ for $t < t'$). Obviously such behavior doesn't make any physical sense since it violates the principle of *primitive causality*[4], which stipulates that an effect should never precede its cause [44]. In general, amplifying or attenuating any bandwidth of input frequencies will yield an acausal response unless these frequencies are also accompanied by a phase shift [45].

We may impose the condition that the response function is causal by demanding that $G(t-t') = 0$ for $t < t'$. Defining $\tau = t - t'$, the Fourier transform of the response function then goes as

$$G(\omega) = \int_0^\infty \mathrm{d}\tau G(\tau) e^{i\omega\tau}. \tag{2.14}$$

Assuming that $G(\omega)$ vanishes at high frequencies faster than $1/\omega$, we apply the relation [46]

$$\frac{1}{\omega' - \omega} \to \mathcal{P}(\frac{1}{\omega' - \omega}) + i\pi\delta(\omega' - \omega) \tag{2.15}$$

to find that

$$G(\omega) = \frac{1}{i\pi}\mathcal{P}\int_{-\infty}^{\infty} \mathrm{d}\omega' \frac{G(\omega')}{\omega' - \omega}. \tag{2.16}$$

Here $\mathcal{P}$ denotes taking the Cauchy principal value of the integral[5]. Since $G(\omega)$ is

[4] *Primitive causality* will to be distinguished from *Einstein causality*, an idea that will be introduced shortly.

[5] Computing the Cauchy principal value amounts to calculating [47] $\lim_{\epsilon\to\infty} \int_{-\infty}^{\omega-\epsilon} \mathrm{d}\omega' \frac{G(\omega')}{\omega'-\omega} + \int_{\omega+\epsilon}^{\infty} \mathrm{d}\omega' \frac{G(\omega')}{\omega'-\omega}$.

complex, it follows that

$$\begin{aligned} \mathrm{Re}[G(\omega)] &= \frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty} \mathrm{d}\omega' \frac{\mathrm{Im}[G(\omega')]}{\omega' - \omega} \\ \mathrm{Im}[G(\omega)] &= \frac{-1}{\pi}\mathcal{P}\int_{-\infty}^{\infty} \mathrm{d}\omega' \frac{\mathrm{Re}[G(\omega')]}{\omega' - \omega}. \end{aligned} \tag{2.17}$$

In the case of a dielectric medium's response to an imposed electromagnetic field, the incoming stimulus and the resulting response are real[6]. Accordingly, the real and imaginary parts of the response function ($G(\omega)_R$ and $G(\omega)_I$, respectively) must obey $G(-\omega)_R = G(\omega)$ and $G(-\omega)_I = -G(\omega)$. We can then recast Eq. 2.17 in a form that does not integrate over negative frequencies:

$$\begin{aligned} \mathrm{Re}[G(\omega)] &= \frac{2}{\pi}\mathcal{P}\int_{0}^{\infty} \mathrm{d}\omega' \frac{\omega'\mathrm{Im}[G(\omega')]}{\omega'^2 - \omega^2} \\ \mathrm{Im}[G(\omega)] &= \frac{-2\omega}{\pi}\mathcal{P}\int_{0}^{\infty} \mathrm{d}\omega' \frac{\mathrm{Re}[G(\omega')]}{\omega'^2 - \omega^2}. \end{aligned} \tag{2.18}$$

These are the famous Kramers–Kronig relations, and they stipulate exactly how the frequency components of a system's response function must be phase shifted relative to the drive such that the response satisfies causality. For linear dielectrics, these equations are often expressed in terms of the frequency–dependent index of refraction $n(\omega)$, which is relates the polarization $P(\omega)$ of a dielectric to an applied electric field $E(\omega)$ according to $P(\omega) = \epsilon_0(n^2(\omega) - 1)E(\omega)$. For some arbitrary real frequency Ω, a similar application of Cauchy's theorem yields

$$n(\omega) - n(\Omega) = \frac{\omega - z}{i\pi}\mathcal{P}\int_{-\infty}^{\infty} \mathrm{d}\omega' \frac{n(\omega') - n(\Omega)}{(\omega' - z)(\omega' - \omega)} \tag{2.19}$$

[6] Although it is often convenient to express the electromagnetic field using complex algebra, it is a *real* field (otherwise it would give rise to complex forces, which is meaningless).

where we have introduced a small positive z such that the resulting pole lies in the lower half–plane. Upon subtracting the same expression with $\omega = \Omega$ (which just amounts to subtracting nothing since $n(\Omega) - n(\Omega) = 0$),

$$n(\omega) - n(\Omega) = \frac{1}{i\pi}\mathcal{P}\int_{-\infty}^{\infty} \mathrm{d}\omega' \frac{n(\omega') - n(\Omega)}{(\omega' - z)}\left(\frac{\omega - z}{\omega' - \omega} - \frac{\Omega - z}{\omega - \Omega}\right). \tag{2.20}$$

In the limit that $\Omega \to \infty$, the second term in Eq. 2.20 goes to unity and we assume that the dielectric medium can no longer respond ($n(\infty) = 1$). Imposing the same symmetry requirements on the positive and negative frequencies of $n(\omega)$ as before, the integral can be performed only over positive frequencies to yield

$$\mathrm{Re}[n(\omega)] = 1 + \frac{2}{\pi}\mathcal{P}\int_0^{\infty} \mathrm{d}\omega' \frac{\omega' \mathrm{Im}[n(\omega')]}{\omega'^2 - \omega^2} \tag{2.21}$$

$$\mathrm{Im}[n(\omega)] = -\frac{2\omega}{\pi}\mathcal{P}\int_0^{\infty} \mathrm{d}\omega' \frac{(\mathrm{Re}[n(\omega')] - 1)}{\omega'^2 - \omega^2} \tag{2.22}$$

$$\tag{2.23}$$

From Eq. 2.21–2.22, it follows that any sharp variation in the imaginary part of the refractive index begets a sharp variation in the real part of the index. In other words, if some band of an applied stimulus sees a lot of gain or absorption, neighboring frequencies will suffer significant phase shifts in order to preserve causality [44]. Let's take the example of a dilute atomic gas. In the linear response approximation, an electric field $\mathbf{E}(t) = \mathbf{E}\cos\omega t$ induces an electric dipole moment $\mathbf{p}$ given by [48]

$$\mathbf{p} = \alpha(\omega)\mathbf{E}\cos\omega t. \tag{2.24}$$

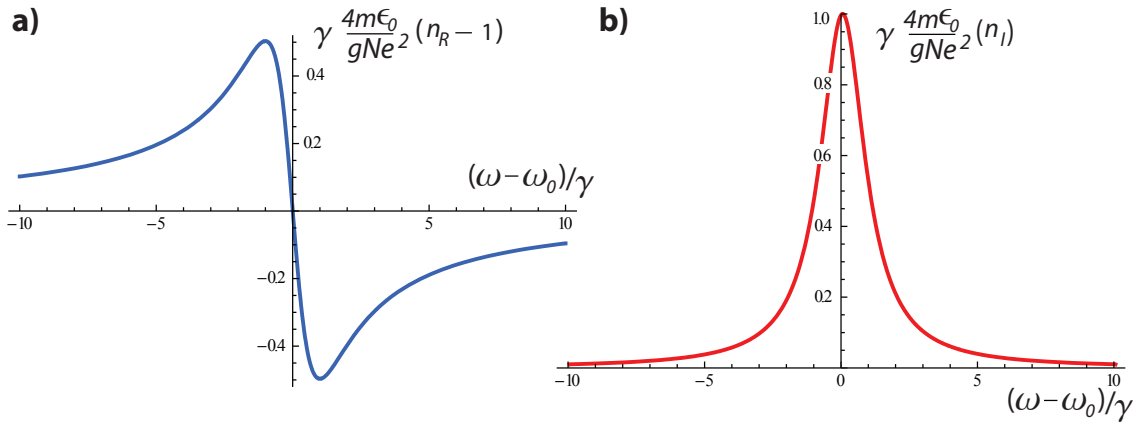


Fig. 2.1: a) Real and b) imaginary parts of the complex refractive index for a dilute absorbing gas. The real and imaginary components are given in terms of the unitless quantity $\frac{Nge^2}{4m\Gamma\epsilon_0}$ where Γ is the decay rate of the excited state, N is the atomic density of the ensemble, and g is the transition oscillator strength. Note how an absorption line yields a steep anomalous dispersion on resonance.

Here $\alpha(\omega)$ denotes the atomic polarizability, and the polarization density $\mathbf{P}$ of the ensemble is given by $\mathbf{P} = N\mathbf{p}$ where N denotes the atomic density of the ensemble. Using the usual constitutive relations for linear dielectric media [39], we find that $n^2(\omega) = 1 + \frac{N}{\epsilon_0}\alpha(\omega)$. A truncated binomial expansion and application of the "Kramers–Heisenberg formula" (obtained using perturbation theory [49]) yields [42]

$$n(\omega) = 1 + \frac{gNe^2}{2m\epsilon_0}\frac{1}{\omega_0^2 - \omega^2 - 2i\gamma\omega}. \tag{2.25}$$

The real and imaginary parts of this complex refractive index are plotted in Fig. 2.1. The figure illustrates how the steepest dispersion occurs in the neighborhood of a resonance ($\omega = \omega_0$), which is consistent with the Kramers–Kronig relations.

2.1.2 *Interpreting the Group Velocity*

From the form of Eq. 2.11, we find that sufficiently steep anomalous dispersion allows v_g to assume negative values, implying that the phase and group velocities of a traveling waveform move in opposite directions. How are such negative group velocities achieved in the laboratory? If a forward–moving pulse initially encounters a medium from the left, how could a leftward moving pulse have any room to propagate unless it starts from the right side of the cell? Remarkably, this is actually what happens [50].

To understand this behavior, let's study the motion of a Gaussian pulse with carrier frequency ω_c that impinges on the front face of fast–light medium at $z = 0$ according to

$$E(z = 0, t) = e^{-i\omega_c t} e^{-t^2/2\tau^2}. \tag{2.26}$$

Here we will take the common approach of expressing the real electric field using a complex representation with the understanding that we will be computing the intensity at the end of the calculation (given by the squared modulus of the field). We will also assume that magnitude of the real part of the refractive index can be well–approximated by unity (which is a very good assumption [51]). Finally, we will ignore the reflection coefficient at the boundaries of the fast–light medium which arise due to the Fresnel equations [52]. Plugging the field $\mathbf{E}_\omega(z, t) = \mathbf{E}_0(z)e^{-i\omega t}$ into the wave equation

$$\nabla^2 \mathbf{E}_0 + n^2(\omega)\frac{\omega^2}{c^2}\mathbf{E}_0, \tag{2.27}$$

we find that each Fourier component obeys $\mathbf{E}_\omega(z,t) = \mathbf{E}_0 e^{i\omega(n(\omega)z/c-t)}$. The full spatiotemporal solution of the propagating wave can be constructed by multiplying the Fourier transform of Eq. 2.26 by the phase factor $e^{i\omega n(\omega)z/c}$ and Fourier transforming back:

$$E(z,t) = \frac{\tau}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}\omega e^{-i\omega t} e^{i\omega n(\omega)z/c} e^{-1/2(\omega-\omega_c)\tau^2}. \tag{2.28}$$

For a pulse whose bandwidth exceeds the linewidth of the atomic resonance, its propagation dynamics can be solved numerically and will generally lead to pulse breakup and distortion [53]. To arrive at some level of intuition about the behavior of the pulse, however, we will follow the discussion developed by Garrett and McCumber [54] and assume that the pulse bandwidth is much smaller than the width of the medium's absorption line. Additionally, their model assumes that the real part of the index of refraction can be approximated as:

$$n(\omega) = 1 - \frac{\omega_0 \omega_p}{\omega(\omega - \omega_0 + i\gamma)} \tag{2.29}$$

where ω_p denotes the medium's plasma frequency [39].

For a sufficiently narrowband pulse, the product of $\omega n(\omega)$ in the exponential term of Eq. 2.28 can be Taylor expanded according to

$$\omega n(\omega) \approx \omega_c n(\omega_c) + (\omega - \omega_c) \underbrace{\left[\partial_\omega(\omega n(\omega))\right]_{\omega_c}}_{\equiv\alpha} + \frac{1}{2}(\omega - \omega_c)^2 \underbrace{\left[\partial_\omega^2(\omega n(\omega))\right]_{\omega_c}}_{\equiv\beta}. \tag{2.30}$$

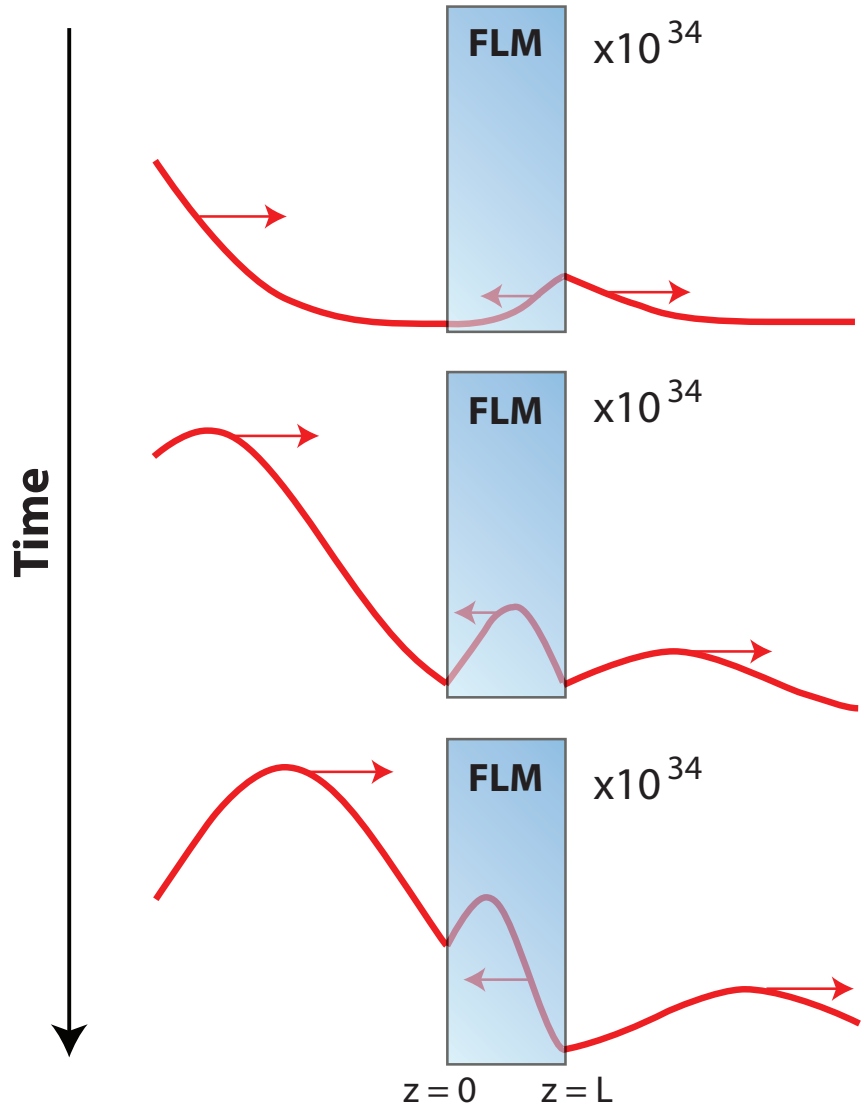


Fig. 2.2: Illustration based on calculations performed by Dmitry Budker [55] of a narrow-band pulse impinging on an absorbing fast–light medium (FLM). Note that the response on right–hand side of the medium has been scaled dramatically by a factor of 10^{34} since a very large absorption coefficient has been assumed (the scaling within the medium varies exponentially across). Negative group velocities *have* been clearly observed in an optical fiber in a beautiful experiment performed by Gehring *et al.* [54].

Evaluating the integral for the electric field then yields [42]

$$E(z,t) \approx \frac{e^{-i\omega_c(t-n(\omega_c)z/c)}}{\sqrt{1-i\beta z/c\tau^2}} \exp\left(\frac{-(t-\alpha z/c)^2}{2\tau^2(1-i\beta z/c\tau^2}\right). \tag{2.31}$$

Assuming that the pulse propagates a sufficiently short distance such that it is not significantly distorted,

$$\frac{i\beta z}{c\tau^2} \ll 1. \tag{2.32}$$

It can be shown that, on resonance ($\omega_c = \omega_0$), the intensity $I(z,t) = |E(z,t)|^2$ of the pulse reduces to

$$I(z,t) = \exp\left(-\frac{2\omega_0\omega_p}{\gamma}\frac{z}{c} - \frac{(t-z/v_g)^2}{\tau^2 - 2\omega_0\omega_p z/\gamma^3 c}\right) \tag{2.33}$$

where $v_g = \frac{c}{1-\omega_0\omega_p/\gamma^2}$. This is a revealing result. The second term of the exponential conveys that the pulse will in general propagate through the medium with a different width than when in vacuum. Additionally, it conveys that the pulse's propagation speed is given by the group velocity v_g.

Since any physical medium is finite in length, the backward propagating electric field originating at the opposite face of the cell ($z = L$ in Fig. 2.2) defines the initial conditions for the forward propagating solution in free space. The field here can be solved using the same impulse response function as in Eq. 2.28 with the exception that, in this case, $n(\omega)$ is equal to unity. Such a frequency–independent refractive index implies that, on the other side of the cell, the outgoing pulse propagates undistorted at a group velocity equal to the speed of light.

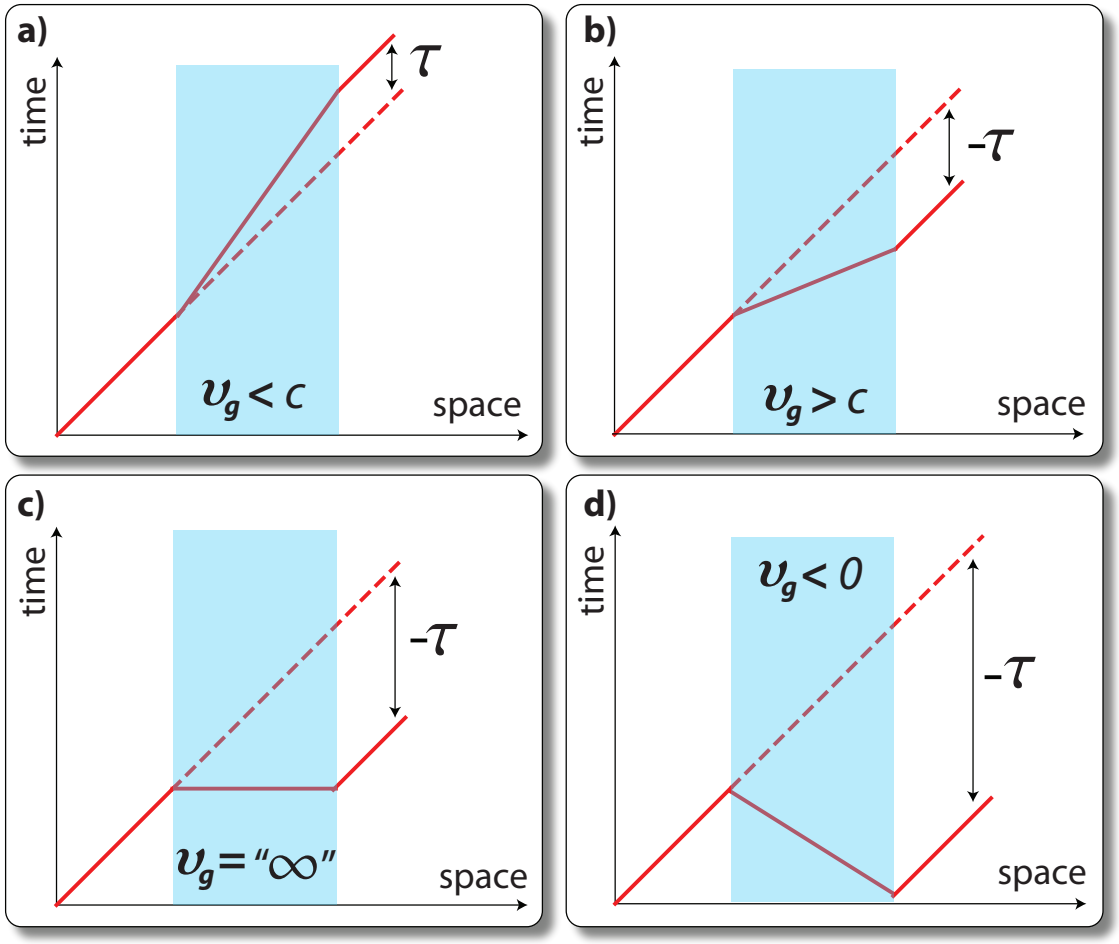


Fig. 2.3: Illustration for recalling the relationship between the magnitude and sign of the group velocity and the corresponding delay or advance. The solid lines represent the motion of the pulse peak when the group velocity is a) slower than c, b) faster than c, c) "infinite", and d) negative. The dotted lines indicate the position of the peak when propagating through vacuum. Notice that the slow–light medium yields a delay in the peak arrival time, while cases b)–d) yield increasingly large advances in the pulse's arrival time.

An interesting interpretational question arises from a cursory examination of Eq. 2.11, which reveals that sufficiently steep anomalous dispersion yields a small negative group velocity. Simultaneously, however, the advance in the arrival time of the peak of the outgoing pulse is expected to increase monotonically with the magnitude of anomalous dispersion. A convenient interpretation of this behavior [56] is illustrated in the Minkowski diagrams of Fig. 2.3 where the speed of light has been normalized to unity[7]. In these figures, small group velocities correspond to steeply sloped lines. Large group delays (or advances), on the other hand, correspond to large deviations from the dotted lines denoting pulse propagation in vacuum. A superluminal group velocity limits the temporal advance of the pulse peak to $\tau < \frac{L}{v_g - c}$ while negative group velocities can achieve advances $\tau > \frac{L}{v_g - c}$. Note that in the case of a negative group velocity, there is a period of time when three peaks exist simultaneously on the inside and outside of the cell.

2.2 Relativity and Information

We have seen that causal response functions allow for superluminal or negative group velocities in media with anomalous dispersion, which in turn give rise to advanced arrival times for the peaks of smooth, slowly–varying pulses. This alone might give the uninitiated reader pause. How can a medium respond to an incident pulse in a causal way and still apparently react before the onset of the stimulus? Something seems to be wrong with causality.

[7] The lines associated with these diagrams should not be confused with information. We will encounter a more careful use of Minkowski diagrams in a later section.

There may appear to be even more subtle offenses to the notion of causality. Intuitively, it might seem reasonable to suppose that the group velocity could be interpreted as the velocity of any information carried by the pulse [57]. If the peak of a pulse arrives earlier at a detector in the presence of a fast–light medium, does this mean that the information content of the pulse also arrived earlier than it would have if the pulse had propagated through vacuum? Practically speaking, it is universally accepted that the answer is no.

The idea that information can only propagate at or below the speed of light lies at the heart of a principle which has come to be termed *Einstein causality*. Einstein causality is a little more subtle than the principle of *primitive causality* that we have already encountered. To understand Einstein causality, it is helpful to recall the notions of the *absolute past* and the *absolute future*.

2.2.1 Absolute Past, Absolute Future

It is now well understood that "simultaneity" is a relative concept [58]. Consider, for example, the x-axis in a given inertial frame's coordinate system. All events that lie on this axis are, by definition, simultaneous in time from that particular frame. In any other inertial observer's frame of reference, however, the x-axis inevitably gets tilted at some angle. Any events lying on a common position axis in one frame, do not lie on the same position axis in any other frame. Simultaneity is completely a frame–dependent concept.

Along the same lines, consider for a moment the locus of events that lie *parallel* to the position axis of some reference frame. These events are also simultaneous in

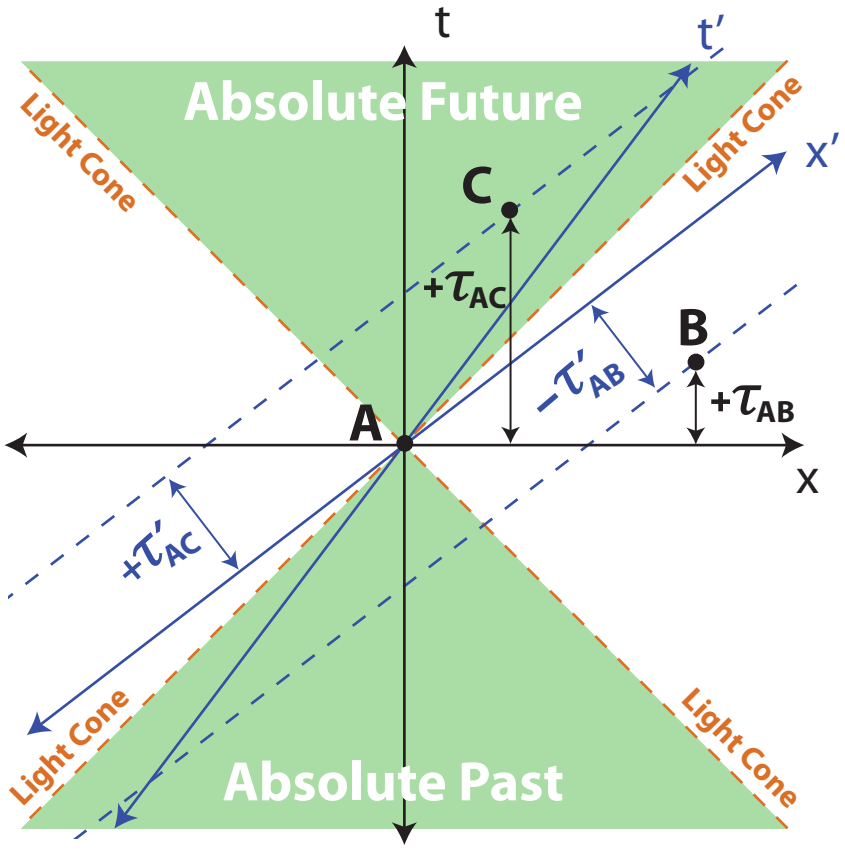


Fig. 2.4: Defining the absolute past and the absolute future by the light cone. All horizontal lines parallel to the x-axis represent contours of simultaneity in time. Similarly, for a moving observer (primed frame), all lines of simultaneity are oriented parallel to the x'-axis (dotted lines). For these lines of simultaneity, one can plainly see that event B occurs after event **A** in the unprimed frame of reference (time difference $+\tau_{AB}$). To a moving observer in the primed frame, event **B** occurs before event **A** with a time difference of $-\tau'_{AB}$). Contrastingly, *all* relativistic observers will agree that event **C** occurs after event **A**–they will just disagree by how much time.

time, and these lines of simultaneity reveal another very important idea. The order in time between two events which lie outside of another's light cone (defined in Fig. 2.4) is also a frame–dependent concept. This can be clearly seen by considering the three events illustrated in Fig. 2.4. In the unprimed coordinate system, event **A** occurs before event **B** by a time difference of τ_{AB}. In the inertial frame defined by the primed coordinate system, however, event **B** occurs *before* event **A** with a time difference of $-\tau'_{AB}$ as can be seen by studying the lines of simultaneity.

For events that lie within one another's light cone, however, it is not possible to find a line of simultaneity that intersects the segment $\overline{AC}$ from below given any relativistically permissible inertial observer. In other words, the light cone establishes the notion of an *absolute past* and an *absolute future* for any given event. For two events located within one another's light cones, all inertial observers will agree which event occurred first (they will just disagree by how much time). Events **A** and **C** in Fig. 2.4 are examples of such events, which are said to be time–like separated (no reference frame exists which can make them simultaneous in time).

We see now that relativity predicts troubling consequences to any hypothetical superluminal transfer of information. If it were possible to send information faster than light, then it would follow that two space–like separated events could be causally connected. For example, imagine that event **A** in the Minkowski diagram in Fig. 2.4 denotes a communication event which causes another event to occur at **B**. In the unprimed frame, there is no apparent violation of *primitive causality* since **B** occurs after **A**. In the primed frame, however, event **B** occurs *before* event **A**.

2.2.2 Where is the Information?

Why then does "fast light" not violate this more subtle relativistic notion of causality? To answer this question, physicists have argued about where the information content of the pulse lives [59]. Consider any pulse that can be described by an analytic function. Given such an analytic function, it is possible to infer the entirety of the pulse's shape by knowing the function's height and all of its derivatives at any point. In light of this observation, it is not obvious why the peak of the pulse should be thought of as the sole location of the pulse's information content. The same information applied to any point in the pulse's tails will similarly do.

So where then does the information reside in an analytic waveform? One popular interpretation actually emerged in the earliest investigations of the subject conducted by Brillouin and Sommerfeld in the early 20^{th} century [60, 61]. I will adopt a slightly modified presentation[8] that still captures the essential physics by considering how a *smooth* waveform with compact support[9] might behave upon encountering a fast–light medium [63]. Consider, for example, the field envelope defined by the piecewise function

$$E(z=0,t) = \exp\left(\frac{-1}{1-(1+t)^2}\right)\Theta(-t)\Theta(t+2) \tag{2.34}$$

[8] The analysis considered by Sommerfeld and Brillouin actually involved studying how a sharp discontinuity in a sine wave propagates through a fast–light medium [42, 62]. I find it interesting to consider what happens when you try to turn the light pulse on as *gently* as possible, avoiding any obvious discontinuity in the derivatives. I'm not sure how fundamental this is, but the fact that it doesn't change the results is interesting.

[9] By compact support, I refer to a traveling wave which is only "turned on" for a finite duration of time and completely "turned off" for all other times.

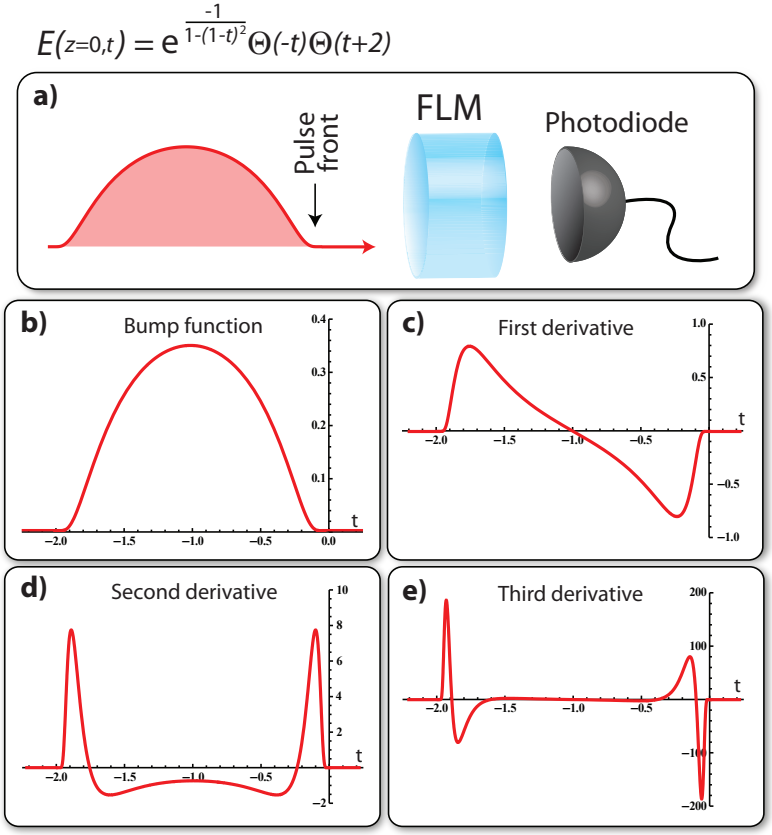


Fig. 2.5: a) The pulse front of a "bump" function (Eq. 2.34) incident on fast–light medium. Classically speaking, a photodiode will not detect any advance in the arrival of the pulse beyond the front (compared to vacuum propagation). The bump and its first three derivatives are given in b)–e) to give a flavor of their behavior. All the derivatives vanish $t \to 0^-$ since they are of the form of a decaying exponential multiplied by a diverging polynomial (the exponential dies faster).

where

$$\Theta(t) = \begin{cases} 0, & \text{for } t \leq 0 \\ \\ 1, & \text{for } t > 0. \end{cases} \tag{2.35}$$

Our piecewise function (referred to as a "bump function," [64]) turns on smoothly in the sense that all of its derivatives continuously vary (see Fig. 2.5). Intuitively, such a pulse seems like a physically reasonable starting point for our consideration.

How would such a pulse behave when encountering a medium with anomalous dispersion? By inserting the Fourier transform of Eq. 2.34 into Eq. 2.28 and assuming that $n(\omega) \to 1$ as $\omega \to \pm\infty$, it can be shown [42, 63] that the propagation integral vanishes for $z > \tau/c$. Moreover, the dynamics of such a pulse incident on a fast–light medium have been numerically simulated [63] and shown to yield strong pulse steepening and even shock–like behavior (Fig. 2.6a–b). When compared to a Gaussian pulse (which is nonzero everywhere in space), the point in the bump function where the pulse initially "turns on" ($t = 0$ in Eq. 2.34) appears to play a crucial role (Fig. 2.6a).

The critical idea to notice here is that this turn–on point is non–analytic (note from Fig. 2.6 that the function and all of its derivatives vanish at $t = 0$, but the function does not vanish everywhere as would be expected by a Taylor series). It is therefore impossible to infer the shape of the pulse until this non–analytic point has passed, i.e. when $t > 0$. This critical point in the field evolution is often referred to as the pulse "front," and it is thought to *always* lead the remainder of the pulse and *always* travel precisely at the speed of light in vacuum [42]. Accordingly, the pulse

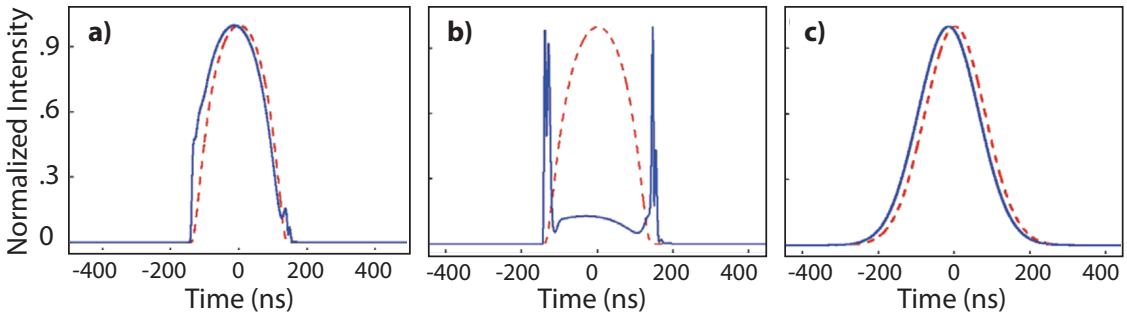


Fig. 2.6: Numerical results reported in [63] for a "bump" function propagating through a fast–light medium obeying the Kramers–Kronig relations. The aspect ratio of the figure has been adjusted only to fit this thesis format. The following caption accompanied the figure: "Temporal behaviour of the [bump function] pulse with nonanalytical points for different propagation distances through vacuum (dashed red line) and fast-light medium (solid blue line): (a) z = 1.2 m and (b) z = 2.4 m. (c) Temporal behaviour of the Gaussian pulse after propagation through vacuum (dashed red line) and fast-light medium (solid blue line) for z = 2.4 m distance."

front is often interpreted as containing the entirety of all "new information" arriving with the traveling waveform. Others have interpreted this idea more broadly by suggesting [59] that the collection of all non–analytic points of a traveling waveform carry the entirety of the pulse's information content. These ideas have received a degree of empirical support in experiments [65] involving bright optical pulses propagating through a vapor of pumped K atoms acting as a fast–light medium. An operational information velocity associated with sharp variations in the amplitude of resonant optical pulses traveling through the vapor was measured to lie below c despite the presence of the fast–light medium.

So far, the discussion has been confined to a purely wave–like description of light fields and (as we shall see in the next chapter) a classically–rooted description of information. Introducing quantum effects to the problem complicates matters, and it is the subject we turn to next.

2.3 Quantum Mechanics and Einstein Causality

2.3.1 Entanglement and No Signaling

It has been frequently observed [66] that information cannot travel superluminally because the carriers of information are physical systems, and physical systems cannot propagate faster than the speed of light [58]. As Landauer famously phrased it [67], "information is physical." But what can we say about truly non–local phenomena such as entanglement? Why can't entanglement be exploited in some clever way to communicate superluminally? For example, consider the collapse of the polarization–entangled state of twin photons prepared according to Fig. 2.7a. The consequences of measuring the polarization of one of these photons is manifestly nonlocal: the two–photon state collapses, and the measurement outcomes display strong correlations that cannot be understood classically. If these measurements are space–like separated, how could this not result in a space–like transfer of information?

To gain some intuition, let's specifically consider what would happen if these entangled photons were sent off in opposite directions to be measured by Alice and Bob (Fig. 2.7b). If Alice measures the polarization of her photon slightly *before* Bob (say in the HV basis yielding the outcome of $|V\rangle$), then the outcome of Bob's measurement will depend on his choice of basis. If Bob measures the polarization of his photon in the HV basis as well, he is guaranteed to get the result $|H\rangle$. On the other hand, if he performs a projective measurement in the right–left (RL) circular

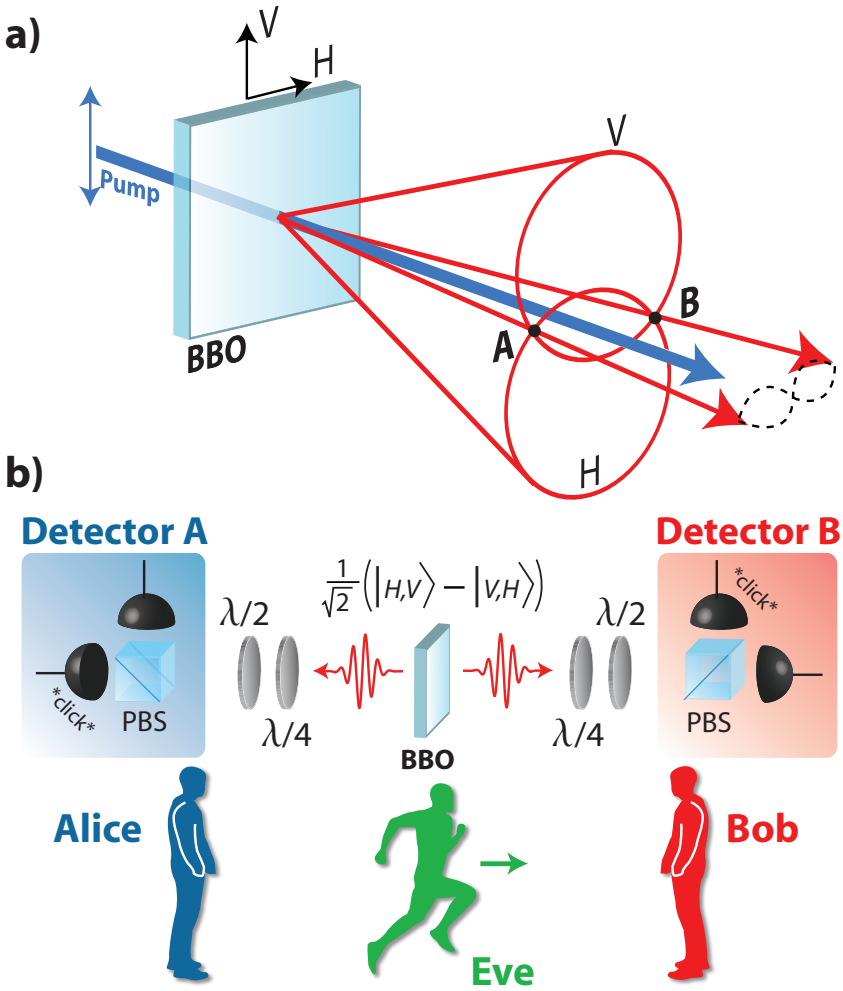


Fig. 2.7: a) Typical experiment to generate a polarization–entangled state $|\psi\rangle = \frac{1}{\sqrt{2}}(|H,V\rangle + e^{i\phi}|V,H\rangle)$ using type–II phase matched parametric down–conversion [68]. The non–linear medium, a BBO crystal, converts a single photon from a strong pump beam into a pair of photons with vertical (V) and horizontal (H) polarizations. At points A and B, there is no "which–path" information to tell the photons apart, so the polarization state is entangled. b) Alice and Bob, who are at rest in the laboratory frame, each detect the polarization of a single photon of the 2–photon entangled state $|\psi\rangle = \frac{1}{\sqrt{2}}(|H,V\rangle - |V,H\rangle)$ with space–like separation. In their frame of reference, Alice detects the polarization of her photon slightly before Bob detects his. Eve, who is moving at a relativistic velocity to the right, will see Bob's detection occurring first. *All* observers see the same correlated measurement outcomes, but the observers might have trouble agreeing on who "collapsed" the state!

polarization basis, his polarization measurement could go either way (which we will denote by the states $|L\rangle$ and $|R\rangle$). It seems then that in performing her measurement first, Alice has done *something* to the state which limits Bob's possible outcomes.

Although Alice cannot actually exert any control over which way her state collapses, the intuition that entanglement can be used to communicate is understandable. Einstein himself puzzled over why such "spooky action at a distance" would not imply the superluminal transfer of information [69]. The alert reader, however, might have noticed a subtle issue lurking in the previous paragraph: we were meant to consider what would happen if Alice measured the polarization of her photon *before* Bob measured his. Any student of relativity should object! After all, for space–like separated detection events (which would be required to communicate superluminally), the notion that Alice or Bob performed the first measurement depends on the chosen inertial frame of reference[10]!

Let's examine the specific case where Alice picks the VH basis for her measurement, Bob picks the RL basis for his measurement, and the following outcome occurs:

	Basis	Outcome
Alice	HV	$\|V\rangle$
Bob	RL	$\|R\rangle$

Observers who see Alice's detector click first would understandably conclude that Bob's state collapses to $|H\rangle = \frac{1}{\sqrt{2}}(|R\rangle + |L\rangle)$. From there, Bob's outcome of $|R\rangle$

[10] I am not discussing how the wave function is affected by Lorentz transformations or anything like that. I am referring to waveplate settings and clicks of a detector, which are macroscopic settings/outcomes of an experiment.

is just the result of further projective state collapse. On the other hand, observers who see Bob's detector click first would conclude that it is *Alice's* polarization state that gets projected into $|L\rangle = \frac{1}{\sqrt{2}}(|H\rangle - i\,|V\rangle)$! Then *Alice's* outcome of $|V\rangle$ will seem like a further projection. Trying to draw conclusions about who collapsed the state or who is messaging whom becomes bewildering.

2.3.2 "Fast Light" with Single Photons

We have yet to confront a rather obvious question: is it possible to advance a single photon using a fast–light medium? This has been the subject of lively debate [70–73], and the answers that have been offered can be quite subtle since a range of issues must be confronted. For instance, since photons are detected probabilistically in time (for a given photon "wave function"), how should any advance be quantified (particularly in the presence of absorption)? Is it possible to define the time at which a photon is created in any meaningful way? How should virtual photons be treated? I will spend this section briefly sketching out a few of these ideas.

Perhaps a good starting point would to be introduce a closely related problem famously taken up by Enrico Fermi in 1932 [74]. Consider an excited two–level atom placed a distance a away from an identical two–level atom in the ground state at $t = 0$ (Fig. 2.8a). The problem asks: what is the probability that the atom in the ground state will be excited at $t > 0$? Fermi originally showed that the excitation probability of atom 2 vanishes outside of the light cone[11] To do so, however, he introduced an approximation (whose justification has been the subject

[11] By this, I mean for times $t < \frac{a}{c}$).

of some debate [75]) without which he would not have obtained this result [42]. Although the calculation is lengthy, a full treatment of the problem predicts [76] an excitation probability for atom 2 that is *exactly causal* (i.e. proportional to $\Theta(t - a/c)$, where $\Theta(x)$ is the Heaviside step function). To my knowledge, the exactly causal solution is generally accepted to be correct. A more relevant question for this thesis is this: what would happen if a dilute gas of 2–level atoms acting as an absorptive fast–light medium were inserted between atoms 1 and 2 (Fig. 2.8b)? Would the presence of the fast–light medium affect the time–dependent excitation probability of atom 2?

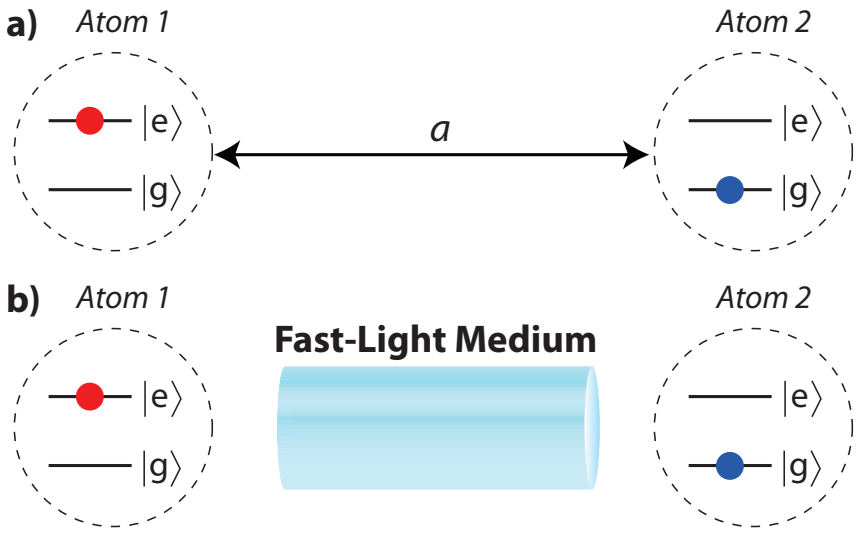


Fig. 2.8: a) Illustration of "Fermi's problem." A pair of two–level atoms are placed a distance a away from one another in vacuum. At $t = 0$ atom 1 is known to be in the excited state and atom 2 is known to be in the ground state. Fermi's problem asks: what is the probability that atom 2 will be excited at time $t > 0$? b) Insertion of a fast–light medium between atoms 1 and 2.

Interestingly, the answer appears to depend on exactly *how* atom 1 is excited [77]. For example, if atom 1 (taken initially in the ground state) were slowly excited

by a weak resonant laser pulse[12], the time–dependent excitation probability $P_2(t)$ of atom 2 can be approximated as [77]

$$P_2(t) \propto e^{-2\omega_0 n_I(\omega_0) z/c} P_1\left(t - \frac{z}{v_g}\right). \tag{2.36}$$

Here $P_1(t)$ denotes the excitation probability for atom 1, n_I the imaginary part of the fast–light medium's refractive index, and ω_0 the resonant frequency of atoms 1 and 2. The group velocity v_g in the argument of $P_1(t)$ corresponds to the same group velocity introduced in Eq. 2.11. So given a steep anomalous dispersion, the peak probability for an excitation of the sink atom (atom 2) occurs *earlier* than the peak excitation probability for the source atom (atom 1). The peak probability is comparatively reduced, however, by the absorptive properties of the atomic gas captured by the imaginary component of the refractive index[13].

On the other hand, in the limit where atom 1 is suddenly excited, the probability of exciting atom 2 outside of the light cone vanishes [77]. It has been suggested that such a sudden excitation plays an analogous role to the pulse front we have already encountered in the classical case. Of course, an infinitely fast excitation of atom 1 should be thought of as just a limiting case and not necessarily physically realistic. Nevertheless, it is still possible to speak in broadly similar terms about single photon advancements as it is with a purely wave–like treatment of fast–light

[12] By "weak," the assumption is that the Rabi frequency is always small compared to the radiative decay rate.

[13] Incidentally, if we tried to get around the absorption problem by putting an actively–pumped *gain* medium between atoms 1 and 2, we'd introduce spontaneously–emitted photons. In this case, we couldn't be sure whether atom 2 was excited by the source atom or from spontaneous emission [42]. This idea will play a central role in the next chapter.

phenomena.

2.4 Quantum Information and Fast Light

This chapter has explored several important ideas to this thesis, particularly the notion that anomalous dispersion cannot be used to send information faster than the speed of light in vacuum. So far we have considered a classically–rooted picture of information, i.e. messages encoded in the controlled modulation of an electromagnetic field in some meaningful sequence. On the other hand, we found that quantum properties like entanglement cannot be used to signal. In this context, it would be understandable to view entanglement as having little value as an information resource. Is this all there is to say about information, or is there be a broader description that can accommodate quantum properties like entanglement?

To begin answering this question, we consider an experiment where Alice tries to send messages to Bob by encoding them in the quantum state of a spin–half particle. Using only the 2 degrees of freedom associated with the particle's spin, Alice can select at most 2 orthogonal states, which we will denote by $\{|\uparrow\rangle, |\downarrow\rangle\}$ in some chosen basis. It is straightforward to see then that Alice can only reliably send 1 bit of information to Bob using the state of her qubit[14].

But what would happen if Bob possessed his own spin–half particle, and Alice and Bob were somehow able to deterministically entangle their particles over an arbitrarily large distance outside of the forward light cone? Since this possibility

[14] For example, to send a "1" ("0"), she might send the state $|\uparrow\rangle$ ($|\downarrow\rangle$). If, on the other hand, she tried to send two bits to Bob by sending one of four distinct states of her spin–half particle, Bob would not be able to perfectly distinguish these states deterministically.

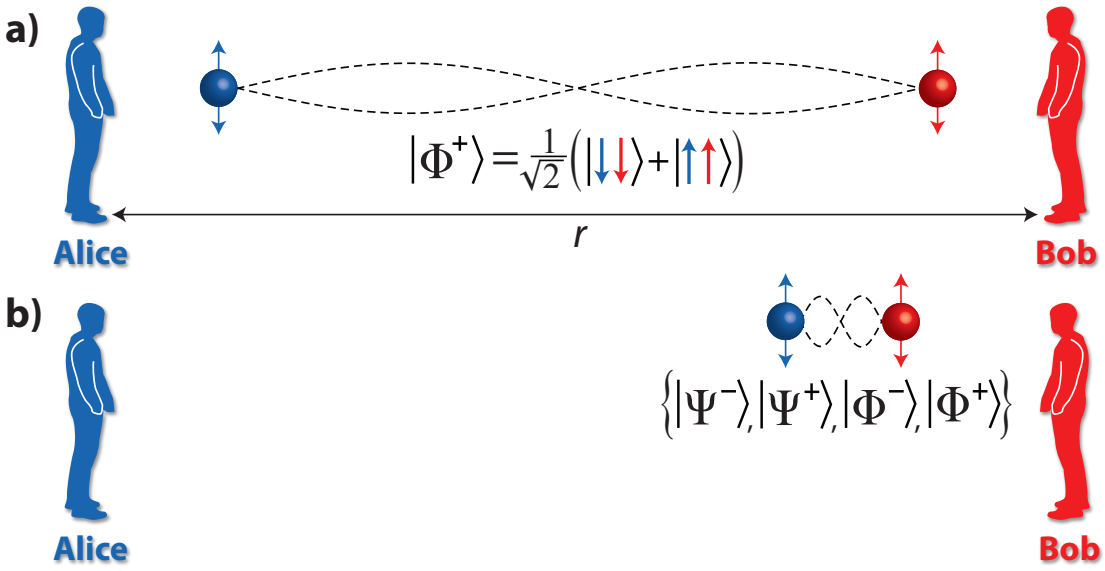


Fig. 2.9: a) Alice and Bob, who are separated by a distance r, share the Bell state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)$. Alice performs a unitary operation (out of the set $\{\hat{I}, \hat{X}, i\hat{Y}, \hat{Z}\}$) on *her qubit alone*, which transforms the shared entangled state into any one of four mutually orthogonal Bell states $\{|\Psi^+\rangle, |\Psi^-\rangle, |\Phi^+\rangle, |\Phi^-\rangle\}$. b) By sending her qubit to Bob, Bob may perform a joint measurement on the qubit pair to uniquely determine the Bell state. Since there are four Bell states to choose from, Alice has sent 2 bits of information along with her single qubit.

does not appear to directly violate Einstein causality (entanglement cannot be used to signal), perhaps there is no immediate reason to object. But would this mean that no information resource of any kind had been established between Alice and Bob? Let's assume Alice and Bob are able to deterministically generate the maximally entangled Bell state

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|\downarrow_A\downarrow_B\rangle + |\uparrow_A\uparrow_B\rangle) \tag{2.37}$$

where subscripts A and B denote the spin–half particle belonging to Alice and Bob, respectively. This state exhibits the interesting property that it is possible to perform a unitary operation on one of the particles which transforms the Bell state

$|\Phi^+\rangle$ into any of the four mutually orthogonal Bell states

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|\downarrow_A\downarrow_B\rangle + |\uparrow_A\uparrow_B\rangle) \tag{2.38}$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|\downarrow_A\uparrow_B\rangle + |\uparrow_A\downarrow_B\rangle) \tag{2.39}$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|\downarrow_A\uparrow_B\rangle - |\uparrow_A\downarrow_B\rangle) \tag{2.40}$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|\downarrow_A\downarrow_B\rangle - |\uparrow_A\uparrow_B\rangle) \tag{2.41}$$

according to

$$\hat{I}_A\,|\Phi^+\rangle = |\Phi^+\rangle \tag{2.42}$$

$$\hat{X}_A\,|\Phi^+\rangle = |\Psi^+\rangle \tag{2.43}$$

$$i\hat{Y}_A\,|\Phi^+\rangle = |\Psi^-\rangle \tag{2.44}$$

$$\hat{Z}_A\,|\Phi^+\rangle = |\Phi^-\rangle\,. \tag{2.45}$$

In Eqs. 2.42–2.45, the operators $\{\hat{X}_A, \hat{Y}_A, \hat{Z}_A\}$ denote the Pauli operators which act on the Hilbert space of Alice's spin–half particle. If Alice performs this operation on her particle and then sends it to Bob, Bob can exploit his possession of both particles to perform a joint measurement of the particle pair which perfectly distinguishes the four Bell states (since the four states are mutually orthogonal). As there are four possible states to choose from, Alice is able to encode 2 bits of information in her choice of the unitary transformation she applies to her particle. Therefore, she effectively sends twice as much information to Bob along with her single qubit than

would have been possible without any shared entanglement present.

Has any kind of information bound been violated in this scenario? In one sense, the answer is no since Alice is sending two bits to Bob using quantum states spanning a four–dimensional Hilbert space. In a four–dimensional space, Alice may select up to four mutually orthogonal states, so her reliable transmission of two bits of information should come as little surprise. In light of this, the name of this protocol ("superdense coding") might even appear to be a misnomer–it merely *saturates* the maximum communicable information [15].

On the other hand, the typical superdense coding protocol involves locally preparing an entangled state before one of the entangled qubits is sent to Alice and the other to Bob. When executed this way, the extra bit of information can be thought of as being sent "off peak" [66] to Bob along with the first spin–half particle. In a sense, the first bit is nascent until Alice decides on her choice of unitary transformation to transform the Bell state. For me, this kind of intuition conforms more closely with the spirit that information is physical and carried by physical systems.

So can a separable state of two particles become entangled outside of the forward light cone? Interestingly, there are claims that this should be possible due to properties of the Feynman propagator [78], though it is unclear to me whether the quality of such entanglement would generally be sufficient to permit useful quantum information protocols. Interestingly, generating maximum entanglement (which is absolutely vital for superdense coding to be possible) between two qubits would

[15] The relevant limit here is the Holevo bound [14], which we will encounter in the next chapter.

require a post–selection process *demanding* that Alice and Bob wait for times $t > \frac{r}{c}$ [78] (where r is defined in Fig. 2.9). I do not know if this condition plays some deeper fundamental role in bounding the behavior of quantum information, but I think it is an interesting question.

We find ourselves confronting one of the central ideas of this thesis: given locally generated entanglement, is it possible to distribute this entanglement outside of the light cone? Surely the consequences of doing so would be important for processes such as teleportation, quantum key distribution, etc. Before treating this question experimentally, we will first take up the task of formalizing the difference between classical and quantum information, which is the subject of the next chapter.

Chapter 3: Information Theory in a (very tiny) Nutshell

"If computers that you build are quantum,

Then spies of all factions will want 'em.

Our codes will all fail,

And they'll read our email,

Till we've crypto that's quantum, and daunt 'em."

– Jennifer and Peter Schor

3.1 Introduction

This chapter seeks to provide a basic and intuitive foundation for understanding the formalism of quantum information theory. This subject will come in handy when we address the question: "how does quantum information associated with entanglement behave when propagating through a fast–light medium with anomalous dispersion?" Our strategy will be to introduce the most relevant ideas in the field of classical information theory first before we continue on to develop many analogous ideas in quantum information.

3.2 Classical Shannon Information

Classical coding and information theory are two very closely connected subjects that not only form the bedrock of modern communication theory but many topics in physics. Essentially, classical information theory seeks to answer two questions:

1. What is the fundamental limit to which data can be compressed?

2. What is the fundamental limit on the rate at which data can be communicated through noisy channels?

The formalism required to answer the first question will allow us to smoothly transition to many important ideas relevant to this thesis. With this in mind, I will devote the first part of this chapter to answering the first of these questions, and I will say just a few words about the second.

3.2.1 The Shannon Entropy

We can quantify the degree of compressibility of an information source by posing (and then answering) the following well defined question: *if Alice wants to communicate a series of messages to Bob, and if she encodes these messages in binary, what is asymptotically*[1] *the shortest average code word that Alice can send to Bob without introducing any errors?* The answer is given by the Shannon entropy. To define the entropy, we treat any given message from the information source as a

[1] By asymptotically, I mean in the limit of infinitely long strings of messages.

random variable X which can assume the message values $X = \{x_1, x_2, ..., x_n\}$. When these messages are independent and identically distributed according to probabilities $\{p_1, p_2, ..., p_r\}$, the entropy is mathematically expressed as [79]

$$H(X) = -\sum_{i=1}^{r} p_i \log_b p_i. \tag{3.1}$$

Here p_i represents the probability of message i from the information source and b denotes the base of the log. For the case of encoding in binary, the appropriate log base is 2 and the entropy is reported in units of bits[2]. Other common choices for the log base (and thus entropy units) are "nats" when $b = e$ or "Hartleys" when $b = 10$. Note that the definition of the entropy only requires the probabilities from the source and not parsing size. In other words, the entropy does not depend on how the output of an information source is encoded (or in what language it is encoded). It is a property of the information source itself.

It can be shown that the average message length of the most efficient binary code (a *Huffman code* [80]) code satisfies [81]

$$H(U) \leq \bar{L} < H(U) + 1 \tag{3.2}$$

where U denotes the random variable representing the possible message outcomes. We see then that information theory bears a very direct relationship with coding theory: the entropy "tells us what is possible" (in terms of the meeting the equality

[2] Unless stated otherwise, in this thesis we will always work in bits and write $\log_2$ simply as log.

$H(U) = \bar{L}$), while coding theory "tells us how to do it" [80].

3.2.2 Interpreting the Entropy

Let's pause now to notice a few important features of the entropy. It can be shown that, for a message alphabet of r possible values, the entropy satisfies [81]

$$0 \leq H(X) \leq \log(r) \tag{3.3}$$

where the equality on the right is satisfied by the probability distribution $\{\frac{1}{r},...,\frac{1}{r}\}$. In other words, the entropy is maximized when the message probabilities are equal. This should make intuitive sense. As the probability of receiving any given message tends towards the probability of receiving any other given message, the information source becomes less and less "predictable" and should require longer codewords on average. Upon observing that

$$\lim_{x \to 1} -x \log x = 0, \tag{3.4}$$

we also see that the entropy vanishes as the probability of receiving one particular message tends to unity.

From the previous discussion, it should be clear that entropy quantifies, in some sense, the "measure" of the message space[3]. As an alphabet of possible messages grows, the message space increases. Similarly, for a given alphabet size, as the

[3] In the literature, the word that is often used is the "uncertainty" of the source. I will avoid this term since we often will deal with Heisenberg uncertainties and quantum information in the same context.

message probabilities become more uniform, the message space increases. In this way, the entropy can be thought to quantify a receiver's (Bob's) degree of ignorance about the next possible message. An equivalent way of viewing this is to think of the entropy is quantifying how much is learned (by Bob) *on average* upon receiving a message.

These interpretations of the entropy closely connect with Ralph Hartley's conception of the "self–information" of a given message [82]. Given a source of random messages with associated probabilities p_i, Hartley quantified the self–information of the message as $-\log(p_i)$. From the definition, it is apparent that rare messages carry more information than more common messages. This also makes sense. For example, contestants on the show *Wheel of Fortune* learn far more about an unknown word when they learn it contains a letter like the letter "z" or the letter "q" than, say, the letter "e." This is because the letters "q" and "z" occur with approximately a 0.1% probability, while the letter "e" has an incidence closer to 13% [81]. The factor of p_i in front of the Log function in Eq. 3.1, however, acts to lower the weight of these large self-information terms on average precisely because they don't occur very often.

3.2.3 Information Processing Systems

An information processing system can be understood as being divided [80, 81] into 5 parts (Fig. 3.1):

1. **Source**: where the information originates.

2. **Encoder**: prepares the information for storage or transmission (which might include adding redundancy to combat the degrading influence of noisy channels).

3. **Channel**: medium of information transportation or storage.

4. **Decoder**: processes the encoded information into its original form (any encoded redundancy is removed).

5. **Information sink**: the final user.

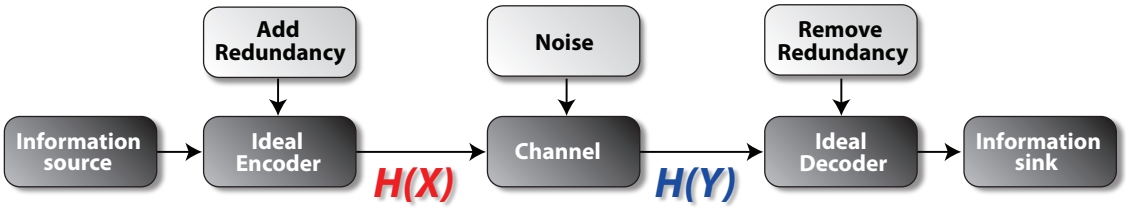


Fig. 3.1: Components of an information processing system. Random variables X and Y describe encoded and decoded messages for the information source and information sink, respectively. In general, $H(X) \neq H(Y)$.

Practically speaking, the limiting step in the faithful transfer of information is modeled by treating the information channel as being noisy. The general approach taken to describe information's travel through the channel is to assign two random variables X and Y to the input and output messages, respectively. The input message X enters the channel, the channel does something to the message (with some probability that is assumed to be known), yielding the output message Y. A noiseless channel yields a one–to–one correspondence between random variables X and Y. Any deviation from such a one–to–one relationship is ultimately what limits

the quality of the information channel. We now turn to the mathematical machinery of information theory appropriate to characterize the degree of correlation between variables X and Y.

3.2.4 Conditional Entropy

Let's begin with the ideal case where the output variable Y exhibits a one–to–one relationship with the input X. In more mathematical terms,

$$P(y_i|x_j) = \delta_{i,j} \tag{3.5}$$

where the conditional probability $p(y_i|x_j)$ has been defined as the probability that outcome y_i will be obtained given the input x_j. This is obviously the ideal situation in any communication scheme: a received message uniquely corresponds to the message that was sent. Of course, the opposite extreme is the situation where variables Y and X are completely unrelated, which can be expressed as

$$P(y_i|x_j) = P(y_i). \tag{3.6}$$

This situation makes communication impossible.

Using the definition of the entropy for a single variable as a guide, we write

down the entropy of the joint input and output space (i.e. the "joint entropy") as

$$\begin{aligned} H(X,Y) &= -\sum_{i,j} P(x_i,y_j)\log(P(x_i,y_j)) \\ &= -\sum_{i,j} P(x_i,y_j)\log(P(x_i)P(y_j|x_i)) \\ &= \underbrace{-\sum_{i} P(x_i)\log(P(x_i))}_{=H(X)} \underbrace{-\sum_{i,j} P(x_i,y_j)\log(P(y_j|x_i))}_{\equiv H(Y|X)}. \end{aligned} \tag{3.7}$$

Here $P_{X,Y}(x_i, y_j)$ is the joint probability distribution, $P(y_j|x_i)$ is the conditional probability distribution, and we have defined $H(Y|X)$ as the conditional entropy[4] [81]:

$$H(Y|X) = -\sum_{i,j} P(x_i,y_j)\log P(y_j|x_i). \tag{3.8}$$

It might be helpful to point out that the third line uses the definition of the marginal distribution $P(x_i) = \sum_j P(x_i, y_j)$. Using the same reasoning, it is easy to show that $H(X,Y) = H(Y) + H(Y|X)$, so

$$\begin{aligned} H(X,Y) &= H(X) + H(Y|X) \\ &= H(Y) + H(X|Y). \end{aligned} \tag{3.9}$$

The interpretation of Eq. 3.9 is instructive: the conditional entropy $H(Y|X)$ quantifies the remaining entropy of Y given that X is known. In more intuitive language, the conditional entropy tells us how much we still don't know about Y given that we have learned the value of X. Accordingly, when the probability

[4] Also referred to as the "equivocation" [80].

distributions for X and Y are independent,

$$\begin{aligned} \underbrace{P(x_i, y_j) = P(x_i)P(y_j)}_{\text{independent } X \text{ and } Y} \rightarrow \;\; & H(X,Y) = H(X) + H(Y) \\ & H(Y|X) = H(Y) \\ & H(X|Y) = H(X). \end{aligned} \tag{3.10}$$

Quite reasonably, the total entropy for two statistically independent distributions is just the sum of the entropies of the distributions. On the other hand, when X and Y are perfectly correlated, each term in the conditional entropy vanishes (taking the convention that $x \log(x) = 0$) and $H(X,Y) = H(X)$. In other words, the "message space" of two perfectly correlated random variables is only as large as the message space of either random variable. Upon learning one of two perfectly correlated random variables, there is no remaining ignorance about what the other random variable could be.

In the interest of making a connection with quantum information theory later, we will rewrite Eq. 3.8 in a rather suggestive form:

$$\begin{aligned} H(Y|X) &= -\sum_{i,j} P(x_i, y_j) \log P(y_j|x_i) \\ &= -\sum_j P(x_j) \sum_i P(y_i|x_j) \log P(y_i|x_j) \\ &\equiv \sum_j P(x_j) H(Y|X = x_j). \end{aligned} \tag{3.11}$$

3.2.5 *Mutual Information*

We are now in a position to introduce the mutual information $I(X;Y)$, which quantifies the gain in information (lowering of the entropy) of X after receiving Y [81]. From our previous encounter with the conditional entropy, we would be justified to suspect that the mutual information and conditional entropy are intimately related[5]. To prove this assertion, let's start with a quantity which is reminiscent of the "self-information" according to Hartley [82]. The mutual information obtained by a single outcome $\{x_i, y_j\}$ can be quantified according to [80]

$$\begin{aligned} I(x_i;y_j) = & \underbrace{-\log(P(x_i))}_{\text{self-information of } x_i} + \underbrace{\log(P(x_i|y_j))}_{\text{entropy after receiving } y_j} \\ & = \log\left(\frac{P(x_i|y_j)}{P(x_i)}\right). \end{aligned} \tag{3.12}$$

The average (or "system") mutual information between the information source and the information sink follows by inserting an extra joint probability factor and taking the sum just as with the entropy:

$$\begin{aligned} I(X;Y) &= \sum_{i,J} P(x_i,y_j) I(x_i;y_j) \\ &= \sum_{i,j} P(x_i,y_j) \log\left(\frac{P(x_i,y_j)}{P(x_i)P(y_j)}\right). \end{aligned} \tag{3.13}$$

Proof of the final line follows similarly to the steps outlined in Eq. 3.9. Note that if X and Y are independent, $P(x_i|y_j) = P(x_i)$ and $I(X;Y) = 0$. At the other extreme, when variables X and Y are perfectly correlated $I(X;Y) = H(X) = H(Y)$.

[5] For classical probabilities! We will run into difficulties with the quantum conditional entropy.

Let me offer two interpretations of the mutual information that might be conceptually helpful. One intuitive way of thinking about mutual information is noting that Eq. 3.13 is an example of the relative entropy $H(p(x)||q(x))$, which can be thought of as a measure of the "distance" between two probability distributions:

$$H(p(x)||q(x)) = \sum_x p(x) \log \frac{p(x)}{q(x)}. \tag{3.14}$$

In this case, the probability distributions being compared are $P(x_i, y_j)$ and $P(x_i)P(y_j)$. As the joint probability distribution $P(x_i, y_j)$ becomes more and more dissimilar to $P(x_i)P(y_j)$ (i.e. X and Y become "less independent" from one another), the distance (mutual information) between these distributions increases. Maximizing the distance between these distributions amounts to maximizing the mutual information, which is accomplished when variables X and Y are perfectly correlated.

A second interpretation of the mutual information follows by recasting Eq. 3.13 in a different form:

$$\begin{aligned} I(X;Y) &= \sum_{i,j} P(x_i, y_j) \log \left(\frac{P(x_i, y_j)}{P(x_i)P(y_j)} \right) \\ &= -\sum_i \log(P(x_i)) - \sum_j \log(P(y_j)) \\ &+ \sum_{i,j} P(x_i, y_j) \log(P(x_i, y_j)) \\ &= H(X) + H(Y) - H(X, Y), \end{aligned} \tag{3.15}$$

Upon invoking the definition of the conditional entropy $H(Y|X)$ in Eq. 3.9, we find

that we can also express the mutual information in terms of the conditional entropy[6]:

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X). \tag{3.16}$$

The interpretation comes from recalling that the conditional entropy quantifies the remaining entropy of a random variable upon learning about another random variable. The mutual information then can be interpreted as how much is learned about random variable X or Y given that the other is known [14]. If X and Y are completely uncorrelated, then $H(X|Y) = H(X)$ and $H(Y|X) = H(Y)$ and the mutual information $I(X;Y) = 0$. If X and Y are one–to–one correlated, then there is no remaining ignorance of either random variable upon learning the outcome of the other. Therefore, $H(X|Y) = H(Y|X) = 0$ and $I(X;Y) = H(X) = H(Y)$.

3.2.6 Graphical Interpretation

Figure 3.2 provides a popular graphical mnemonic [14, 80, 81, 83] to remember the relationships among the definitions we've introduced so far in Eq. 3.16. Imagine that out of a universe U of possible message outcomes we are interested in the subsets identified in the Venn diagram of the figure. The size of the subsets are associated with the measures of the message outcome space for each random variable. Again, a larger message space does not necessarily correspond to a larger number of possible messages but a more *randomly distributed* space of messages. In the case of Fig. 3.2, the measure of possible outcomes Y is larger than the input X, which could result

[6] The quantum analogues of Eqs. 3.13 and 3.16 will not hold for bipartite quantum systems, giving rise to a property called *quantum discord*.

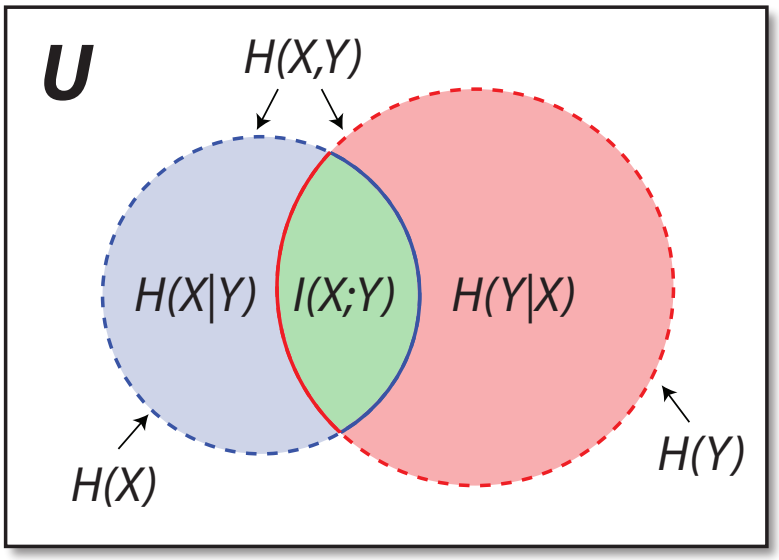


Fig. 3.2: Set theory interpretation of mutual information. The "universe" of possible message outcomes is given by the box labeled U. The total joint entropy $H(X, Y)$ is delineated by the dashed lines, the entropies of the individual random variables $H(X)$ and $H(Y)$ by the red and blue (solid+dashed) lines, the conditional entropies $H(X|Y)$ and $H(X|Y)$ by the blue and red regions, and the mutual information $I(X; Y)$ by the green.

from noise added to the communication channel (for example). With this diagram in mind, the interpretation of Eq. 3.15 becomes more intuitive. The last line of the equation can be interpreted as subtracting the singly counted mutual information of the joint entropy $H(X, Y)$ from the doubly counted mutual information from the sum of $H(X)$ and $H(Y)$.

3.3 Quantum Information

So much for classical information theory. Although we are finally in a position to discuss many important ideas in quantum Shannon theory, it probably comes as little surprise that extending classical Shannon theory to quantum systems can be subtle. To see that classically correlated random variables will, in general, behave

differently than quantum bipartite systems, one need look no further than the example of Bell inequalities. There are even simpler problems fundamentally related to the non–commutativity of observables, however, that make defining the entropy of single qubits a little tricky.

For example, consider the particular qubit state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle). \tag{3.17}$$

If we make projective measurements of this spin state, we will obtain with equal likelihood the outcome of spin up or spin down. Does that mean that the entropy of such a state is 1 bit?

It might be worth considering the following question: what would happen if we performed a series of projective measurements on an ensemble of such qubit states? If we express the eigenstate $|\rightarrow\rangle$ of the $\hat{S}_x$ operator in the z–basis, we obtain

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle). \tag{3.18}$$

Projective measurements of the x–component of the spin yield only one answer, $|\rightarrow\rangle$. This would intuitively suggest an entropy of zero! If we think about the problem a little more generally, we realize that any arbitrary state $|\psi\rangle = \alpha\,|\uparrow\rangle + \beta\,|\downarrow\rangle$ is an eigenstate of the basis aligned with the state's Bloch vector. It looks like we can always make the entropy of any *pure* state look like it's zero if we choose the right basis.

On the other hand, what can we say about the maximally mixed (or maximally "unpolarized" [84]) state[7] $\rho = \frac{1}{2}\mathbb{1}$? If we perform unitary transformations on the state to express ρ in any choice of basis, we obtain $\frac{1}{2}\mathbb{1}$ again. This result is unavoidable because the state's Bloch vector is zero. In other words, it consists of an equal incoherent mixture of states $|\uparrow\rangle$ and $|\downarrow\rangle$ in *any basis*:

$$\rho = \tfrac{1}{2}(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|). \tag{3.19}$$

For this state, we really are *completely ignorant* about what the outcome of any projective measurement will be before we make it. Given our intuition for entropy of classical random variables, it really *does* make sense to assign an entropy of 1 bit to this state.

Another way to view the preceding discussion is as follows: imagine Alice has been given a message encoded in binary and that she wants to send the message to Bob using a stream of spin–half particles. To simplify things, we'll assume that zeros and ones are equally likely (and that Bob knows this ahead of time), yielding a classical entropy of 1 bit. It would make sense for Alice and Bob to agree on a set of *orthogonal* states to encode the zeros and ones of the message since they can be perfectly distinguished. For example, Alice could prepare the state $|\downarrow\rangle$ whenever she wishes to send a zero and the state $|\uparrow\rangle$ whenever she wishes to send a one. Bob could then measure the state of these qubits in the $\{\uparrow,\downarrow\}$ basis, and perfectly decode Alice's message. From Bob's point of view, before he performs a projective

[7] In this chapter, I will suppress the hats when expressing the density matrix ($\hat{\rho} = \rho$) to unclutter the notation a little.

measurement of any particular qubit, he is completely uncertain as to whether it will project up or down. To Bob then, the state of any qubit he receives is $\rho = \frac{1}{2}\mathbb{1}$, and the von Neumann entropy of the state ρ will yield 1 bit, matching the classical entropy of the message.

On the other hand, if Alice sent the same state $|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$ *regardless* of whether she wished to encode a one or zero, it should intuitively be obvious that she would not not be sending any information at all. How could she be sending information if she just sends the same quantum state over and over? Bob's projective measurements in the $\{\uparrow, \downarrow\}$ basis will be completely random, and the outcomes will have nothing to do with what Alice wants to communicate[8].

3.3.1 von Neumann Entropy

States lying somewhere between being totally pure and maximally mixed demand a consistent, flexible treatment. For such a treatment, we need a mathematical operation that is basis–independent and yields a unique scalar value for a state with any degree of mixedness. The trace satisfies these criteria, and it is the big idea behind quantifying the entropy of quantum states[9] [85]:

$$S(\rho) \equiv -\mathrm{Tr}(\rho \log \rho). \tag{3.20}$$

[8] This observation captures the essence of an extremely important idea in quantum information theory, namely the *Holevo Bound* [14].

[9] To avoid confusion, we will give the von Neumann entropy its own notation S to distinguish it from the classical Shannon entropy H.

$S(\rho)$ is referred to as the von Neumann entropy, and it serves as the quantum counterpart for the Shannon entropy in information theory. If λ_i are the eigenvalues of the state ρ, then $S(\rho) = -\sum_i \lambda_i \log(\lambda_i)$. For the maximally mixed spin state given above, we find that indeed $S(\rho) = 1$ bit as we suspected. Similarly, for any pure state the diagonalized density matrix consists of a single eigenvalue equal to unity, which implies an entropy equal to zero (which conforms to our previous discussion).

Our goal now will be to explore how this quantum formulation of the entropy compares to that for classical random variables. Many properties of the von Neumann entropy[10] that will be of interest to us will involve composite quantum systems (in this thesis, we will only consider bipartite states). To understand many of these properties, however, it is very helpful to first introduce a few more ideas first.

3.3.2 Joint Entropy

Extending the definition of the von Neumann entropy to all classes of bipartite states is straightforward [83]:

$$S(\rho_{AB}) = -\operatorname{Tr}(\rho_{AB} \log \rho_{AB}). \tag{3.21}$$

We are now prepared for our first encounter with a major difference between quantum and classical Shannon theory. Recall that the diagonalized density matrix for any pure state consists of a single non-zero eigenvalue equal to unity, implying an

[10] I will just refer to the von Neumann entropy as "the entropy" when it is clear from context that I am referring to quantum systems.

entropy of zero. By characterizing the joint entropy of a pure bipartite state $|\psi_{AB}\rangle$ using the Schmidt decomposition[11], however, we can see that the entropies of the subsystems are both equal and non–zero if $|\psi_{AB}\rangle$ is entangled. To summarize,

$$\begin{aligned} |\psi_{AB}\rangle \quad \rightarrow \qquad & S(\rho_{AB}) = 0 \\ & S(\rho_A) = S(\rho_B) \geq 0. \end{aligned}$$

On the other hand, inspection of Eq. 3.9 reveals that for classical probabilities,

$$\begin{aligned} H(X,Y) &\geq H(X) \\ H(X,Y) &\geq H(Y). \end{aligned} \tag{3.22}$$

The classical and quantum joint entropies can behave quite differently!

3.3.3 Mutual Information

Recall that the classical mutual information $I(X;Y)$ quantifies the gain of information about X given Y. The mutual information is the standard measure of correlations in classical random variables, and this is no different quantum mechanically. In analogy with the final line of Eq. 3.15, we define the quantum mutual information between systems A and B according to

$$I(\rho_B;\rho_A) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \tag{3.23}$$

We should pause for a moment to observe something quite remarkable about

[11] As a reminder, the Schmidt decomposition is a bipartite representation of a composite pure state $|\psi\rangle$ that is always guaranteed to exist (provided that the composite state is pure). Specifically, any composite pure state $|\psi_{AB}\rangle$ can be expressed in terms of orthonormal states $\{|i_a\rangle\}$ and $\{|i_b\rangle\}$ in spaces A and B according to $|\psi_{AB}\rangle \sum_i \lambda_i |i_a\rangle |i_b\rangle$.

Eq. 3.23. Recall that for two *perfectly correlated* classical random variables X and Y, $I(X;Y) = H(X) = H(Y)$. So for example, if $H(X) = H(Y) = 1$ bit, the highest mutual information that can be reached for such a classical system is $I(X;Y) = 1$ bit. In contrast, consider the example of the maximally entangled singlet state[12] $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$. It is a simple matter to confirm that $S(\rho_{AB}) = 0$ due to its purity, and that $S(\rho_A) = S(\rho_B) = 1$ bit. Accordingly, the mutual information of this state $I(\rho_A;\rho_B) = 2$ bits! Here is an important example of the fundamental dichotomy between quantum and classical information theory. The von Neumann entropy has very nicely captured the idea that quantum mechanics permits *stronger* correlations than are allowed classically.

Despite this important difference, the von Neumann mutual information also quantifies correlations between quantum systems that can be understood as being classical. For example, consider what would happen if someone attempted to "spoof" the quantum correlations of the Bell state[13]

$$|\Phi^+\rangle = \frac{1}{2}(|00\rangle + |11\rangle) \tag{3.24}$$

by randomly preparing the states $|00\rangle$ or $|11\rangle$ with an equal probabilities:

[12] Actually, this argument applies to any four of the Bell states, which are all maximally entangled.

[13] To conform more closely with typical conventions in the literature, I am going to start expressing states in "computational basis" notation. This simply amounts to making the substitutions $|\downarrow\rangle \to |0\rangle$ and $|\uparrow\rangle \to |1\rangle$, for example.

State	Probability
$\|00\rangle$	1/2
$\|11\rangle$	1/2

What is the mutual information of such a state? It is straightforward to confirm that the state described by this preparation is given by

$$\rho = \frac{1}{2} |00\rangle \langle 00| + \frac{1}{2} |11\rangle \langle 11| \tag{3.25}$$

and that the mutual information is 1 bit. In fact, this makes sense since the states $|0\rangle$ and $|1\rangle$ are in principle perfectly distinguishable (if measured in the computational basis) and can in a sense be thought of as outcomes of a classical random variable. For two maximally correlated random variables, if one of the variables has two equally likely outcomes, the mutual information is 1 bit. We can therefore understand the mutual information of 1 bit for this state as a quantifying the state's classical correlations. It should now be clear that the von Neumann mutual information quantifies both classical and quantum correlations [86].

3.3.4 A Quantum Conditional Entropy?

Upon inspection of Eq. 3.16, it might seem initially seem to be a simple matter to define the quantum mutual information in terms of a quantum conditional entropy. After some thought, however, it becomes apparent that defining the conditional entropy for quantum systems leads to difficulties (to the best of my knowledge, there is no generally accepted definition). For example, consider the generic bipar-

tite state ρ_{AB}. Upon performing a projective measurement $\Pi_A = |i_A\rangle \langle i_A|$ of system A, one obtains the state [14]

$$\rho_i = \frac{1}{p_i(i)} \Big((|i_A\rangle \langle i_A| \otimes \mathbb{1}_B) \rho_{AB} (|i_A\rangle \langle i_A| \otimes \mathbb{1}_B) \Big), \tag{3.26}$$

where

$$p_i(i) = \mathrm{Tr} \Big((|i_A\rangle \langle i_A| \otimes \mathbb{1}_B) \rho_{AB} \Big). \tag{3.27}$$

It might then seem intuitive to define the conditional entropy according to[14] [83]

$$S(\rho_B | \Pi_A) = \sum_i p_i(i) S(\rho_i), \tag{3.28}$$

which very suggestively resembles Eq. 3.11. On the one hand, this definition seems nice because it directly connects to what we can control in the laboratory. In other words, when we are given a correlated composite system AB and we want to learn something about B without directly measuring it, we can make a measurement of A and exploit the correlations. The difficulty with Eq. 3.28, however, is that the result will generally depend on the *measurement basis*[15]. Although this makes Eq. 3.28 somewhat cumbersome to work with theoretically, it does not present any fundamentally insurmountable problems.

[14] Actually, this definition is more generally extended to POVMs rather than just projective measurements. Since I will not spend time defining POVMs, however, I won't pursue the consequences here.

[15] Unless the state is invariant under all unitary transformations.

3.3.5 Quantum Discord

What would happen if we evaluated Eq. 3.28 given the choice of projective measurement Π_A that *minimized* the resulting entropy of system B? In this way, we would learn as much as possible about system B through a measurement of A. Although in practice this means performing really difficult calculations for larger systems (every possible projective measurement must be considered!), it provides a useful measurement–based quantifier of correlation between two quantum systems.

It might be apparent by now, however, that something strange is going on. The definition of mutual information in Eq. 3.23 does not depend on any choice of measurement since it is evaluated using only the *state* of the composite system. On the other hand, if we try to write down a definition of the mutual information that invokes the conditional entropy analogous to the classical case in Eq. 3.16, the result apparently depends on the specific measurement we elect to make. This is obviously very different from classical information–a classical bit is what it is, and it does not care how you look at it!

Let's press on anyway by using Eq. 3.23 to formulate an alternative interpretation of the mutual information, which we will label with a new symbol, J [87]:

$$J(\rho_B;\Pi_A)_{\{\Pi_A\}} \equiv \max_{\{\Pi_A\}} S(\rho_B) - S(\rho_B|\Pi_A)_{\{\Pi_A\}}. \tag{3.29}$$

The subscript $\{\Pi_A\}$ is meant to convey that $J(\rho_B;\Pi_A)_{\{\Pi_A\}}$ has been evaluated after considering all choices of projective measurements $\{\Pi_A\}$. Note also that we have

chosen the measurement which *maximizes* $J(\rho_B;\Pi_A)_{\{\Pi_A\}}$ so that we obtain as much information about B from a measurement of A as possible.

If we define $\delta(\rho_B;\Pi_A)$ as the difference between Eqs. 3.23 and 3.29

$$\delta(\rho_B;\Pi_A) \equiv I(\rho_B;\rho_A) - J(\rho_B;\Pi_A)_{\{\Pi_A\}}, \tag{3.30}$$

we can confront the following interesting question: is it always possible to pick Π_A such that our two definitions of the mutual information yield the same result (i.e. $\delta(\rho_B;\Pi_A) = 0$)? The answer is no [87]. Consequently, $\delta(\rho_B;\Pi_A)$ has been dubbed the *quantum discord* due to this apparent mismatch between classical and quantum mutual information.

Perhaps the existence of the discord is not terribly surprising. As the example of the mutual information of a Bell state (hopefully) illustrated, analogous ideas in quantum and classical information theory can behave quite differently for entangled states. It should come as little surprise then that entangled states always carry nonzero discord. What may be surprising, however, is that a non–vanishing discord can exist even for completely separable states! Even more bizarrely, realizing nonzero discord for separable states isn't particularly challenging. In fact, just the opposite! Not only is discord present in just any quantum state picked at random from a Hilbert space of any dimension, but it is in general *hard* to keep a correlated state with zero discord from developing a non–zero discord over time [88]! Discord, it seems, is everywhere (including the EPR states discussed in this thesis [89]).

3.4 Information Measures for Gaussian States

Although much of the preceding introduction to quantum information centered mostly on qubit states, conceptually speaking, the definitions are general in scope. Since this thesis deals with continuous–variable entanglement, we would do well to round out the discussion by fleshing out how several of these information measures can be computed using phase–space observables [16]. Generally speaking, this is not an easy task. Fortunately, Gaussian states exhibit a great deal of well–understood symmetry [90], which makes it possible to take significant measurement shortcuts without sacrificing any knowledge about the state. Moreover, since Gaussian states are easily produced in the laboratory, these shortcuts are of great practical importance. It should be easy to see why quantum protocols involving Gaussian states represent a very powerful alternative to qubit protocols.

3.4.1 Entropy of a Single Gaussian Mode

Before developing the machinery needed evaluate the mutual information of a bipartite Gaussian state, let's first describe how the entropy of a *single* Gaussian mode can be evaluated using the statistics of its phase space observables. Proceeding in this manner is ideal for two reasons. For one, it will provide us with two pieces of the expression for the mutual information (Eq. 3.23) that we need to compute

[16] Caution: many of the equations in this section vary significantly in the literature based on normalization conventions for the covariance matrix. These normalization conventions are not unimportant–failing to account for them correctly can lead to completely unphysical results. Although the discussion of this chapter relies heavily on [86], I am using a different normalization convention for the covariance matrix in which the two–mode shot noise is normalized to unity. It just strikes me as much tidier not to carry around factors $\sqrt{2}$, 2, etc.

anyway, so we do not lose anything by starting with the single–mode case. More importantly, however, the two–mode case proceeds by essentially the same kind of arguments.

We saw in Chapter 2 that any Gaussian state can be completely characterized by its covariance matrix. Accordingly, it makes sense to expect that it should be possible to arrive at an analytical relationship between the covariance matrix and the von Neumann entropy. How do we get there? We already saw in Chapter 2 that an ideal vacuum–seeded phase-insensitive (PIA) amplifier yields a state whose quadrature fluctuations are quantified by the variance

$$\langle \sigma^2 \hat{X} \rangle_{\mathrm{PIA}} = \langle \sigma^2 \hat{Y} \rangle_{\mathrm{PIA}} = \cosh 2r. \tag{3.31}$$

This observation, combined with the definition of the two–mode covariance matrix, leads us to conclude that the covariance matrix of the individual modes of an ideal two–mode squeezed state is given by

$$\boldsymbol{\gamma}_{a,b} = \mathrm{Tr}_{b,a}[\boldsymbol{\gamma}] = \begin{pmatrix} \langle \sigma^2 \hat{X} \rangle & 0 \\ 0 & \langle \sigma^2 \hat{Y} \rangle \end{pmatrix}. \tag{3.32}$$

How might we find that entropy of this state? We have already seen that the entropy of a pure state is zero, so it would be intuitive to expect that the entropy of a single mode can be written in terms of its purity, μ. To find such an expression, we exploit the high degree of symmetry of a Gaussian state. For example, *any*

non-displaced[17] single mode Gaussian state can be written as

$$\rho = \hat{S}(\xi)\rho_{\rm th}\hat{S}^{\dagger}(\xi) \tag{3.33}$$

where $\hat{S}(\xi)$ is the squeezing operator and ρ_{th} is the single–mode thermal state (both of which we encountered in the first chapter). Since the squeezing operator $\hat{S}(\xi)$ is unitary, it does not affect the purity

$$\begin{aligned}
\mu &= {\rm Tr}[\rho^2] && (3.34)\\
&= {\rm Tr}[\hat{S}(\xi)\rho_{\rm th}\hat{S}^{\dagger}(\xi)\hat{S}(\xi)\rho_{\rm th}\hat{S}^{\dagger}(\xi)] && (3.35)\\
&= {\rm Tr}[\hat{S}^{\dagger}(\xi)\hat{S}(\xi)\rho_{\rm th}\rho_{\rm th}] && (3.36)\\
&= {\rm Tr}[\rho_{\rm th}^2], && (3.37)
\end{aligned}$$

where we have used the cyclic property of the trace. In other words, the entropy of any single mode Gaussian state is *solely determined* by its thermal component.

Recall from Chapter 2 that any undisplaced thermal state can be uniquely identified by its average photon number $\bar{n}$. While not short, it is relatively straightforward to show [91] that the entropy of a thermal state can also be expressed in terms of $\bar{n}$:

$$S(\rho_{\rm th}) = \bar{n}\log\left(\frac{\bar{n}+1}{\bar{n}}\right) + \log(\bar{n}+1). \tag{3.38}$$

On the other hand, the average photon number is very simply related to the purity

[17] The non-displaced assumption does not affect the generality of this argument since the displacement operator $\hat{D}(\alpha)$ is unitary and does not affect the purity of the state [36].

μ according to

$$\mu = \mathrm{Tr}[\rho_{th}^2] \tag{3.39}$$

$$= \left(\frac{1}{1+\bar{n}}\right)^2 \sum_m \left(\frac{\bar{n}}{1+\bar{n}}\right)^{2m} \tag{3.40}$$

$$= \frac{1}{1+2\bar{n}} \tag{3.41}$$

which allows us to express the von Neumann entropy in terms of μ

$$S(\rho_{\mathrm{th}}) = \frac{1-\mu}{2\mu} \log\left(\frac{1+\mu}{1-\mu}\right) - \log\left(\frac{2\mu}{1+\mu}\right). \tag{3.42}$$

From the discussion of the Wigner function in Chapter 2 (and its relationship with the covariance matrix), we know that the state's purity can be simply computed from the Wigner function

$$\mu = \int_{-\infty}^{\infty} W^2(X,P) dX dY = \frac{1}{\sqrt{\mathrm{Det}(\boldsymbol{\gamma})}}. \tag{3.43}$$

This allows us to express the entropy of *any* single–mode Gaussian state ρ in terms of its covariance matrix. Upon simplification, the entropy can be expressed as [86]

$$S(\rho) = f\left(\frac{1}{2}\sqrt{\mathrm{Det}(\boldsymbol{\gamma})}\right), \tag{3.44}$$

where the function $f(x)$ is defined according to

$$f(x) = (x+\tfrac{1}{2})\log(x+\tfrac{1}{2}) - (x-\tfrac{1}{2})\log(x-\tfrac{1}{2}). \tag{3.45}$$

A plot of $f(x)$ is included in Fig. 3.3.

Equation 3.44 is a profound result! By measuring the variances and covariances of *any* single–mode Gaussian state using homodyne detection, we are able to completely measure the state's von Neumann entropy (which is otherwise not an easy problem).

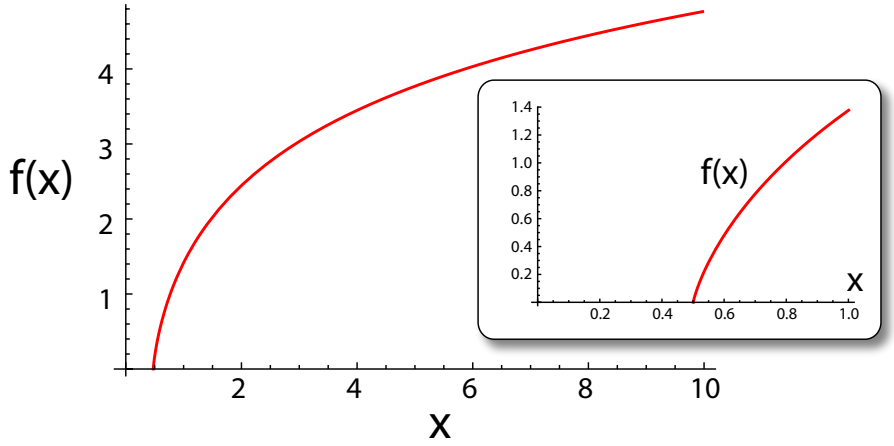


Fig. 3.3: Plot of $f(x) = (x + \frac{1}{2})\log(x + \frac{1}{2}) - (x - \frac{1}{2})\log(x - \frac{1}{2})$. Note that $f(x)$ vanishes for $x = \frac{1}{2}$, which corresponds to the case of a pure state.

3.4.2 Joint Entropy of a Bipartite Gaussian State

The same arguments relevant to the single–mode case can be applied to evaluating the joint entropy of a bipartite state. Any bipartite Gaussian state can be completely characterized by its covariance matrix, and can be expressed in terms of a tensor product of two thermal states, a two–mode squeeze operator, two single mode squeeze operators, and two rotation operators [86]. Although the math involved is more complicated, there is essentially no new physics here: thermal components of the two–mode state solely determine the joint entropy. For the sake of

brevity, I will proceed by simply stating the results (the interested reader may refer to [86] for details).

For an undisplaced[18] bipartite Gaussian state ρ_{ab}, the covariance matrix $\boldsymbol{\gamma}_{ab}$ takes the form

$$\boldsymbol{\gamma}_{ab} = \frac{1}{2}\begin{pmatrix} \langle \hat{X}_a^2 \rangle & 0 & \frac{\langle \hat{X}_a\hat{X}_b + \hat{X}_b\hat{X}_a \rangle}{2} & 0 \\ 0 & \langle \hat{Y}_a^2 \rangle & 0 & \frac{\langle \hat{Y}_a\hat{Y}_b + \hat{Y}_b\hat{Y}_a \rangle}{2} \\ \frac{\langle \hat{X}_a\hat{X}_b + \hat{X}_b\hat{X}_a \rangle}{2} & 0 & \langle \hat{X}_b^2 \rangle & 0 \\ 0 & \frac{\langle \hat{Y}_a\hat{Y}_b + \hat{Y}_b\hat{Y}_a \rangle}{2} & 0 & \langle \hat{Y}_b^2 \rangle \end{pmatrix}. \tag{3.46}$$

For the remaining discussion of this section, it will be convenient to introduce 3 submatrices:

$$\boldsymbol{\alpha} = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}, \tag{3.47}$$

$$\boldsymbol{\beta} = \begin{pmatrix} \gamma_{33} & \gamma_{34} \\ \gamma_{43} & \gamma_{44} \end{pmatrix}, \tag{3.48}$$

$$\boldsymbol{\omega} = \begin{pmatrix} \gamma_{13} & \gamma_{14} \\ \gamma_{23} & \gamma_{24} \end{pmatrix}. \tag{3.49}$$

where γ_{ij} denotes the entry in the i^{th} row and j^{th} column of $\boldsymbol{\gamma}_{ab}$(see Fig. 3.4). From these matrices, we define the quantity

$$\Delta = \text{Det}(\boldsymbol{\alpha}) + \text{Det}(\boldsymbol{\beta}) + 2\text{Det}(\boldsymbol{\omega}). \tag{3.50}$$

[18] By an undisplaced state, I refer to a state whose modes a and b satisfy $\langle \hat{X}_{a,b} \rangle = \langle \hat{Y}_{a,b} \rangle = 0$.

Obtaining Δ in turn allows us to compute related quantities n_+ and n_-:

$$n_\pm = \sqrt{\frac{\Delta \pm \sqrt{\Delta^2 - 4\mathrm{Det}(\boldsymbol{\gamma})}}{8}}. \tag{3.51}$$

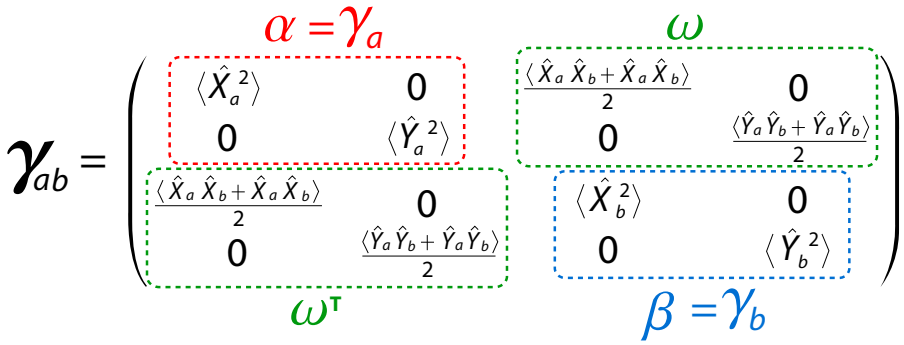


Fig. 3.4: Relevant submatrices of the covariance matrix (when expressed in standard form). $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ represent the covariance matrices describing individual modes a and b while $\boldsymbol{\omega}$ captures the covariances. Note that when $\boldsymbol{\gamma}_{ab}$ is not expressed in standard form, the lower left submatrix is in general equal to $\boldsymbol{\omega}^{\mathrm{T}}$.

The values of $n_\pm$ are referred to as the *symplectic eigenvalues*[19] and uniquely determine the entropy of the two–mode state ρ_{ab}:

$$S(\rho_{ab}) = f(n_-) + f(n_+). \tag{3.52}$$

Here $f(x)$ represents the same function we have already introduced in Eq. 3.45.

Let us take a moment to check that this definition is reasonable. For example, if we consider a pure two–mode squeezed state, we should expect a joint von Neumann entropy of zero. Is this true? In chapter 2, we found that a pure bipartite EPR state

[19] The origins for the term "symplectic" are related to a symmetry imposed on the covariance matrix by the Bosonic commutation relations [90].

is described by the covariance matrix (given my choice of normalization convention)

$$\boldsymbol{\gamma} = \begin{pmatrix} \cosh 2r & 0 & \sinh 2r & 0 \\ 0 & \cosh 2r & 0 & -\sinh 2r \\ \sinh 2r & 0 & \cosh 2r & 0 \\ 0 & -\sinh 2r & 0 & \cosh 2r \end{pmatrix}. \tag{3.53}$$

Computing the determinant of this covariance matrix yields $\mathrm{Det}(\boldsymbol{\gamma})$=1 and a $\Delta = 2$. Accordingly, the $n_{\pm}$ terms both evaluate to $\frac{1}{2}$, which yields a two–mode entropy of zero (just as we expect).

3.4.3 Gaussian Mutual Information

We have now assembled all the pieces necessary to compute the mutual information $I(\rho_a;\rho_b)$ of a bipartite Gaussian state. Expressing the result in terms of sub–matrices defined in Eqs. 3.47–3.49, we conclude that

$$I(\rho_a;\rho_b) = f\left(\frac{1}{2}\sqrt{\mathrm{Det}(\boldsymbol{\alpha})}\right) + f\left(\frac{1}{2}\sqrt{\mathrm{Det}(\boldsymbol{\beta})}\right) - f(n_-) - f(n_+). \tag{3.54}$$

We are therefore able to completely measure the mutual information of any bipartite Gaussian state using statistics obtained by homodyne detection performed at only a few local oscillator phases. We will exploit this convenient fact in the next chapter to study how an entangled state behaves in the presence of anomalous dispersion.

Chapter 4: Quantum Mutual Information of an Entangled State Propagating through a Fast–Light Medium

4.1 Introduction

Over the past decade, many experiments have demonstrated the ability to manipulate the group velocities of optical pulses moving through atomic vapors [92–95]. In particular, much work has been done to understand fast–light phenomena associated with anomalous dispersion, which gives rise to group velocities that are greater than the speed of light in vacuum (or even negative) [96]. For classical pulses propagating without the presence of noise, it has been well established theoretically [60, 61] that the initial turn–on point of a pulse (the "pulse front") propagates through a linear, causal medium at the speed of light in vacuum. It is often argued [59] that this signal carries the entirety of the pulse's classical information content since the remainder of the pulse can in principle be inferred by measuring the pulse height and its derivatives just after the point of non-analyticity has passed.

Experimentally, particularly in the inevitable presence of quantum noise, pulse fronts may not convey the full story of what is readily observed in the laboratory. It is thus interesting to consider other operational definitions of a signal that apply to

particular systems. For example, Stenner, et al. [65] studied the propagation of classical information encoded in bright, actively–shaped optical pulses traveling through a fast–light medium. These experiments showed that the operational information velocity is actually slowed to speeds less than c. Although noise may have affected the experimental results, these experiments were not conducted in a regime where quantum noise necessarily played a crucial role. On the other hand, adopting a definition of signal velocity based on observing a given signal–to–noise ratio, Kuzmich, et al. showed how quantum noise associated with gain–assisted fast–light would be expected to limit the early detection of smooth, narrowband pulses consisting of only a few photons [97].

4.1.1 Description of the Experiment

Throughout this chapter we adopt an alternative definition of a signal by choosing it to be the random, but strongly correlated quantum fluctuations between two spatially–separated parts of a bipartite entangled state. The entangled state in this experiment is generated via four-wave mixing (4WM) in a vapor of ^{85}Rb heated to a temperature $T = 110$ °C [98], which converts two photons from a strong pump beam into "twin" photons emitted into spatially separated modes referred to as the probe and the conjugate (Fig. 4.1a). The fluctuations of the probe and conjugate electric fields are not externally imposed, and they present no obvious pulse fronts or non–analytic features to point to as defining the signal velocity. As such, classically–rooted approaches to defining the signal or information content of the individual modes are not readily applicable to this system.

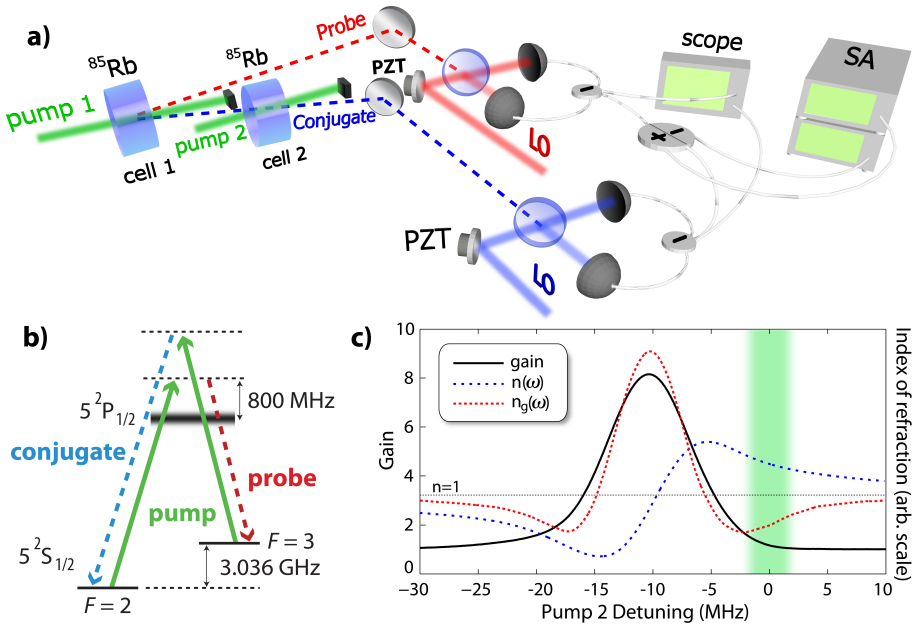


Fig. 4.1: a) Experimental set-up. Vacuum-squeezed twin beams are generated in cell 1 using 4WM in a double-lambda configuration (pictured in part b). We create a region of anomalous dispersion for the conjugate in a second vapor cell using a second 4WM process driven by pump 2, whose frequency is independently tunable with respect to pump 1 (see section 4.4). We scan the phase of the local oscillators (LOs) using piezo-electric transducers (PZT) in order to verify the presence of entanglement. The sum and difference signals of the homodyne detections are recorded on a pair of spectrum analyzers (SAs) to detect quantum correlations. An oscilloscope is triggered to record time traces of the individual homodyne detectors when a predetermined threshold of squeezing is seen on the spectrum analyzers. c) Measured gain profile (black solid line) of the second 4WM process as a function of the detuning of pump 2 relative to pump 1. From the gain profile, we numerically compute the associated refractive index $n(\omega)$ and group index n_g. It should be noted that the scales for the index $n(\omega)$ and group index n_g are different from one another. In determining the advancement, we confine our attention to fluctuations in the frequency band (shaded green) where we observe quantum correlations generated in cell 1. We tune the second pump frequency so that the bandwidth of anomalous dispersion coincides with the bandwidth where we observe quantum correlations.

Despite the randomness of these fluctuations, however, there is quantum information shared between the modes due to the entanglement. Although entanglement cannot be used to signal superluminally [99], it is thought to be an essential resource in quantum information science [90, 100]. Accordingly, the prospect of storing [101] or delaying [102] entanglement has attracted significant interest. To our knowledge, the behavior of entanglement upon propagation through fast–light media has not previously been characterized experimentally. Here we investigate the behavior of the mutual information of an entangled state of vacuum-squeezed twin beams using the formalism of quantum information. We study how the anomalous dispersion associated with phase–insensitive gain [103] affects these correlations by inserting a second vapor cell into the path of the conjugate and driving a second 4WM process with a separate pump (Fig. 4.1b). We show that when one portion of the state passes through this fast–light medium, the peak of the quantum mutual information between the modes is advanced, but the arrival of the leading tail is not. We also show that, in contrast, the leading and trailing tails of the mutual information are both delayed when one of the modes propagates through a slow–light medium with the same level of added noise.

4.1.2 Kramers–Kronig Relations as a Guide

The real and imaginary parts of the nonlinear susceptibility $\chi^{(3)}$ that govern the response of the second 4WM process to the conjugate can be described by a set of equations similar to the Kramers-Kronig relations applicable to linear dielectric media [104]. These relations stipulate that the gain profile of the medium will

generally give rise to a frequency band of anomalous dispersion [103] on the wings of the gain line. Using this relationship between gain and dispersion as a guide (see Fig. 4.1c), we change the detuning of the pump beam used to drive the fast–light 4WM process such that the conjugate frequency overlaps with the region of anomalous dispersion. Since the frequency of the conjugate beam is approximately 4 GHz to the blue of the atomic absorption line, blocking the second pump renders the atomic vapor in the second cell nearly dispersionless and nonabsorptive. Thus, comparing transmission of the conjugate through the cell with the pump unblocked or blocked permits a comparison between propagation through a fast–light medium or free space, respectively.

4.2 Experimental Results

By performing separate balanced homodyne detections of the probe and conjugate modes, we measure the fluctuations of the in-phase ($\hat{X}$) and out-of-phase ($\hat{Y}$) amplitudes of the electromagnetic field in each beam, which are referred to as the field quadratures. Taken individually, the probe and conjugate beams exhibit quadrature fluctuations that exceed the shot–noise limit. Taken together, however, these fluctuations display strong correlations beyond the limits achievable classically. To characterize the strength of the correlations, it is helpful to introduce the joint quadrature operators $\hat{X}_- = (\hat{X}_p - \hat{X}_c)/\sqrt{2}$ and $\hat{Y}_+ = (\hat{Y}_p + \hat{Y}_c)/\sqrt{2}$ where p and c denote the probe and conjugate fields, respectively. For the appropriate choice of local oscillator phases [98], the fluctuations of one of the joint quadratures ($\langle \Delta \hat{X}_-^2 \rangle$

or $\langle \Delta \hat{Y}_+^2 \rangle$) fall below the shot–noise limit (are "squeezed").

4.2.1 The Inseparability Criterion

We verify the presence of entanglement by calculating a related quantity, the inseparability $\mathcal{I}$ (recall the discussion in Chapter 1):

$$\mathcal{I} \equiv \langle \Delta \hat{X}_-^2 \rangle_m + \langle \Delta \hat{Y}_+^2 \rangle_m. \tag{4.1}$$

Here $\langle \Delta \hat{X}_-^2 \rangle_m$ is the minimum value of the difference signal, $\langle \Delta \hat{Y}_+^2 \rangle_m$ is the minimum value of the sum, and each term is normalized to the shot–noise limit. An inseparability $\mathcal{I} < 2$ is a necessary and sufficient condition to conclude that any bipartite Gaussian state is entangled [16]. We measure the inseparability by allowing the local oscillator phases to drift while continuously triggering an oscilloscope to record time traces of each individual homodyne detector. We postselect time traces that are recorded when the local oscillator phases allow us to observe the squeezing in one joint quadrature (refer to section 4.4 for details). We then add or subtract the homodyne time traces (according to which joint quadrature is squeezed), Fourier transform the result, and integrate the power spectral density over a 100 kHz–2 MHz bandwidth. By acquiring noise power statistics for both joint quadratures, we calculate the inseparability.

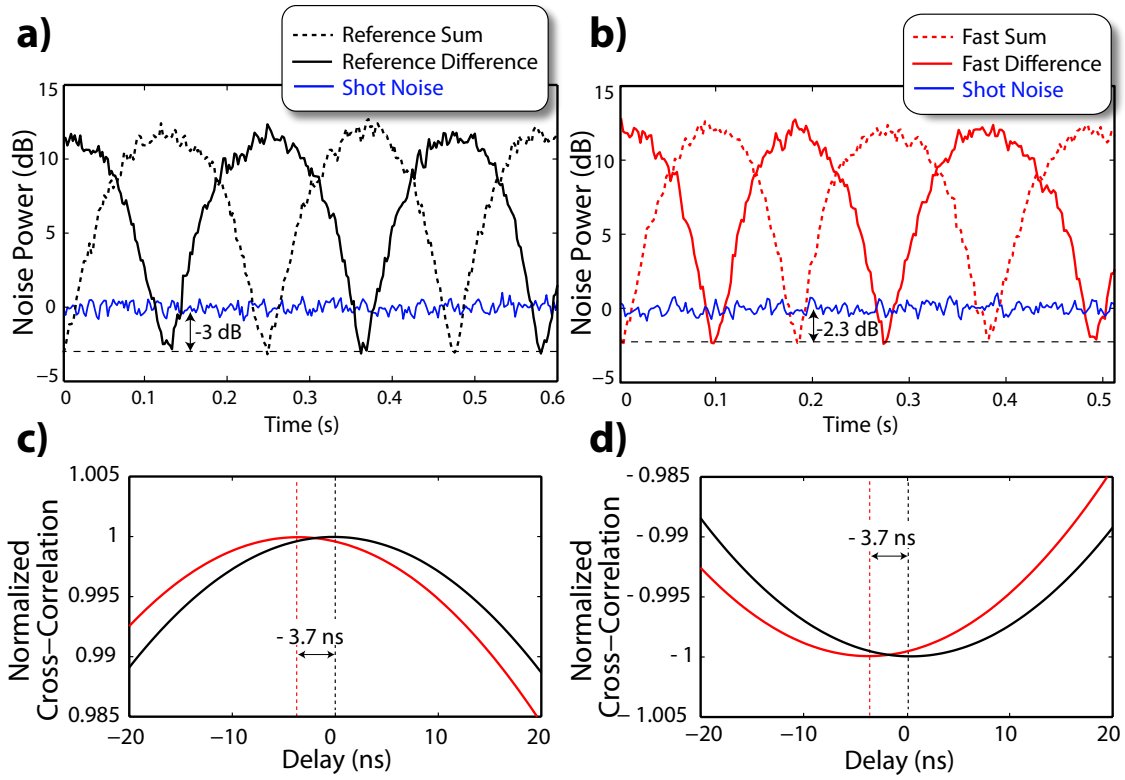


Fig. 4.2: Persistence of correlations associated with entanglement in the presence of anomalous dispersion. a) We observe up to -3 dB of squeezing with an associated inseparability $\mathcal{I} \approx 1$ when the second (fast–light) four-wave mixing process is suppressed. b) In the presence of a small phase–insensitive gain giving rise to anomalous dispersion, the squeezing reduces to -2.3 dB and $\mathcal{I}$ increases to 1.2, which is still sufficient to show entanglement ($\mathcal{I} < 2$). c) Average normalized cross-correlation functions for the correlated and d) anti-correlated joint quadratures. The reference and fast–light data are both subject to a low-pass filter (see section 4.4) used to match the frequency band of the cross-correlation function to that used to calculate $\mathcal{I}$.

4.2.2 Trade–Off between Noise and Advancement

Studies of bright beam propagation through fast–light media [105] have investigated the trade–off between the magnitude of advancement and the amount of added noise [29, 106] as a function of detuning. Here we choose a detuning of the second pump that produces a readily detectable advancement of the conjugate fluctuations without significantly deteriorating the inseparability. By operating in a regime of low gain (G $\approx$ 1.1), we maintained an inseparability of $\mathcal{I} = 1.2$ under fast–light conditions, confirming the persistence of entanglement between the probe and conjugate after the conjugate passes through the fast–light medium (Fig. 4.2a).

Other experiments involving bright classical pulses propagating through 4WM–based fast–light media showed that lower seed intensities lead to smaller advancements [107], making it unclear a priori if it would be possible to detect the advancement of a few–photon state of light. A standard practice to demonstrate an advance or delay in the arrival of a pulse peak traveling through a dispersive medium is to scale the maximum of the output pulse to match that of the input (since dispersion is typically accompanied by gain or loss). Analogously, we confirm that the fluctuations of the continuous–wave conjugate are advanced through the fast–light cell by computing the normalized cross–correlation function of the detected probe and conjugate quadratures for both the reference and fast–light cases (Fig. 4.2c–e). We evaluate the correlation functions after filtering the probe and conjugate homodyne time traces with a 100 kHz–2 MHz bandpass filter (see section 4.4) to insure that the cross-correlation only uses fluctuations in the frequency range that was used to

evaluate the inseparability, $\mathcal{I}$. After acquiring 200 time traces, we conclude that the peak of the cross-correlation function is shifted forward in time by 3.7 ± 0.1 ns, corresponding to a fractional advance of $\approx 1\%$ relative to the cross-correlation width (approximately 300 ns). Here, the uncertainty is estimated by taking the standard deviation of the mean for the cross–correlation peak advancements over all the experiments.

While useful to clearly see an advancement of the correlations, the normalized cross–correlation function of the field quadratures does not capture how the noise added through phase–insensitive gain of the fast–light 4WM process affects the transport of any quantum information. As highlighted by Kuzmich, et al. [97], understanding the effects of added noise is critical since they could play a vital role in limiting effective signal speeds. We studied the strength of the entanglement as a function of the relative delay by plotting (Fig. 4.3a) the average inseparability, $\mathcal{I}$, as a function of the relative delay (see section 4.4). The delay is implemented in software in exactly the same manner as when calculating the cross–correlation function. Although the location of the minimum value of $\mathcal{I}$ is advanced in time for the fast–light case, the degradation of $\mathcal{I}$ acts, within the experimental uncertainty, to prevent the leading tail from advancing forward in time. Figure 4.3b provides a sampling of the delay–dependent squeezing measurements used to calculate the inseparability, which indicates an advance in the maximum squeezing of 3.7 ± 0.1 ns as one would expect (Fig. 4.3c).

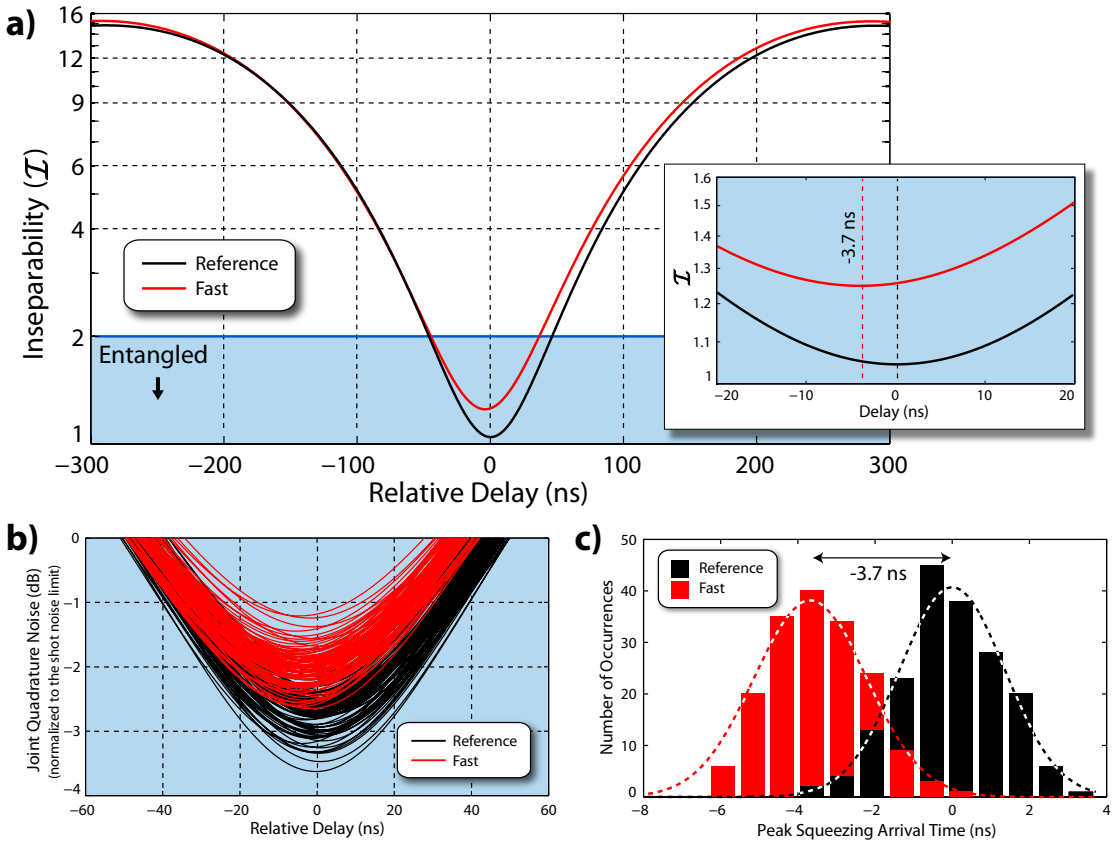


Fig. 4.3: a) Average advance and accompanying degradation of the inseparability, $\mathcal{I}$ ($\mathcal{I} < 2$ implies entanglement), in the presence of anomalous dispersion (fast, red curves) and upon blocking the second pump (reference, black curves). b) Sampling of the squeezing versus delay over 200 experimental iterations used to compute the average $\mathcal{I}$. c) Histogram of the sampled minima of the joint quadrature noise (i.e. maximum squeezing) versus the relative probe–conjugate delay. We fit the sampled shots to a Gaussian distribution to extract an advance of 3.7 ± 0.1 ns where the uncertainty has been estimated by computing the standard deviation of the mean. The reduced χ^2 of the fast and reference distributions were 0.91 and 0.68, respectively. These χ^2 values respectively yield P values of 0.52 and 0.73.

4.2.3 von Neumann Mutual Information

In our experiment, where we are measuring the continuous random fluctuations of the probe and conjugate beams, there is no imposed "signal" as such. The fluctuations on one beam, however, carry information about the fluctuations on the other. We can capture this by calculating the quantum mutual information between the two beams (Fig. 4.4), working from the same basic data as used to calculate the delay–dependent inseparability. The mutual information quantifies the total (classical plus quantum) correlations between the probe and conjugate [87]. To calculate the mutual information of our twin beam state, we exploit the fact that any bipartite Gaussian state can be completely characterized by the variances and covariances of the field quadratures [86]. By assuming that the state of the twin beams is Gaussian, we are able to calculate the mutual information by retrieving relevant statistical moments of the quadrature fluctuations from the photocurrent time traces [89]. In good agreement with the squeezing and cross-correlation measurements, we observe an advancement of 3.7 ± 0.1 ns of the peak of the delay–dependent mutual information, paired with a degradation due to uncorrelated noise added by the fast–light cell (see section 4.4). This degradation appears to prevent us from observing an advance of the leading tail of the fast–light mutual information (red curve in Fig. 4.4).

While superluminal information velocities would violate Einstein causality, subluminal information velocities do not. This observation has led to the misconception that slow group velocities might necessarily limit the propagation speed of classical information in slow–light media. In a follow–up study involving classical

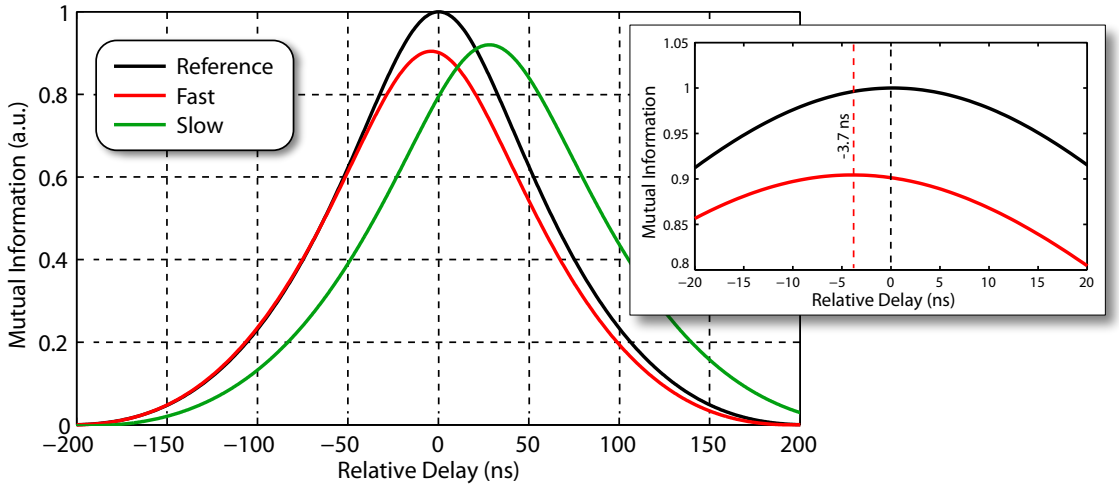


Fig. 4.4: Comparison of the computed quantum mutual information between the probe and conjugate as a function of the relative delay for the cases of fast and slow light. When considering the fast–light advancement of the conjugate (red trace), we observe an advance in the peak of the mutual information of 3.7 ± 0.1 ns. There is no statistically significant advance of the leading tail of the mutual information in the case of fast–light propagation. Repeating the same analysis for slow–light propagation of the probe, we observe a delay of both the leading and trailing tail of the mutual information by hundreds of standard deviations (green trace).

pulses in slow–light media, Stenner, et al. found [108] that the velocity of classical information should be viewed as distinct from the group velocity for slow–light pulses. Moreover, it has been suggested [59, 109] that the set of "non–analytic" points in physical waveforms should be thought of as the only carriers of classical information, which by bandwidth arguments must travel through the medium precisely at c. Under this interpretation, dispersive media in general do not affect the propagation of classical information.

4.2.4 Experiments with Slow Light

Given these classically-rooted interpretations, it might seem plausible to expect that if the fast–light medium in our experiment were replaced with a slow–light

medium, the delayed mutual information would lie within the envelope defined by the reference case (no delay). Using the techniques outlined in [102], we tuned the temperature and the pump detuning of the fast–light 4WM process to slow the propagation of the probe. We slowed the probe to the greatest extent possible while limiting the degradation of the inseparability to the same level as with the fast–light case. The behavior of the delay–dependent slow–light mutual information (green trace in Fig. 4.4) is plotted alongside the reference and fast–light cases. Given an equivalent degradation of the quantum mutual information with added noise, we are able to observe significant delays of the leading and trailing tails of the mutual information compared to the reference case. Additionally, we observe broadening of the full-width at half-maximum of the delayed mutual information, which is generally consistent with the character of distortion expected in slow–light systems [110]. These results highlight behavioral differences between fast and slow light in the quantum regime.

4.3 Conclusions

To summarize, we have demonstrated an advancement in time of the quantum fluctuations in one mode of an entangled state of light passing through a fast–light medium while preserving the entanglement between the modes. We showed that the peak of the quantum mutual information between the modes can be advanced in time, but added noise associated with the dispersion prevents us from observing an advance of the leading tail. In contrast, in a slow–light medium operating under

conditions which produce a similar reduction in the peak mutual information, the leading and trailing tails of the mutual information can be significantly delayed. We hope that this work will motivate further inquiry into the propagation of quantum information through fast and slow–light media.

4.4 Technical Details

Here we develop in more detail the techniques used to reproducibly achieve the same dispersive properties in the fast–light four–wave mixing process over the course of many measurements, a necessary condition for gathering statistics. We also provide further details regarding the detection settings and data analysis used to show an advance in the quantum correlations. We briefly describe the data required to compute the quantum (von Neumann) mutual information, which is based on [86, 89]. Finally, we conclude by presenting a simple theoretical model showing how the inseparability parameter $\mathcal{I}$ and von Neumann mutual information $I(P;C)$ are affected by the introduction of a phase–insensitive gain to one of the twin beams.

4.4.1 Laser System

In order to reproducibly manipulate the dispersive properties of the ^{85}Rb vapor in the fast–light vapor cell, it was essential that the small relative detuning (tens of MHz) between the two pumps be stable and easily tunable to within ≈ 100 kHz. To achieve these requirements, we used two double-passed acousto-optic modulators

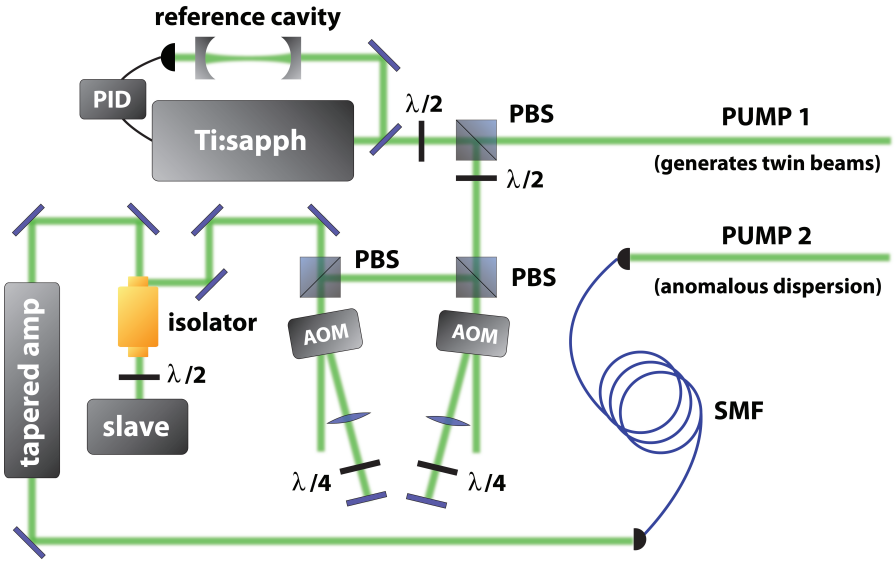


Fig. 4.5: Preparation of the fast–light pump. 10 mW of power is picked off from a Ti:sapphire laser (frequency–stabilized to a reference cavity) using a half–wave plate ($\lambda/2$) and a polarizing beamsplitter (PBS). This light was then coupled into two acousto-optic modulators (AOMs) arranged in oppositely–shifted double–pass configurations with small difference in their frequencies. The output of the second AOM was used to injection lock a slave diode laser, and the slave output was amplified using a tapered amplifier. The spatial mode of this shifted, amplified light was spatially filtered using a single-mode fiber (SMF) before being used to drive the second (fast–light) four-wave mixing process.

(AOMs) to shift the frequency of approximately 10 mW of light picked off from a master Ti:sapphire laser, which itself was stabilized to a reference cavity (Fig. 4.5). The cascaded double-passed AOMs enabled us to tune the second pump beam relative to the first by up to ±100 MHz. Double–passing the AOMs was necessary in order to decouple the shift in frequency from the direction of propagation so that the shifted light could be used to injection–lock a slave diode.

The output of the slave diode laser was amplified using a tapered amplifier and spatially filtered using single–mode fiber (SMF). Seeding the slave diode with approximately 50 μW of light shifted by the AOMs ensured sufficient optical power ($\approx$ 30 mW) to saturate the tapered amplifier and maintain a stable pump power over the tuning range. After amplification and spatial filtering, we were left with 300 mW of optical power. The mode was gently focused to a waist of approximately 1.5 mm inside of the second cell (which was heated to 105 °C) used to create the region of anomalous dispersion. Once the lock was established, we could tune the master Ti:sapphire laser several GHz without the slave diode losing the lock or observing any change in optical power of the fast–light pump.

4.4.2 General Comments about Data Acquisition

In order to acquire time traces of the two homodyne detections, we allowed the phases of the local oscillators to drift while we synchronously triggered two spectrum analyzers and an oscilloscope. The spectrum analyzers detected the noise power present on the sum and difference signals of the homodyne detections while the oscilloscope was configured to measure the direct time–dependent output of

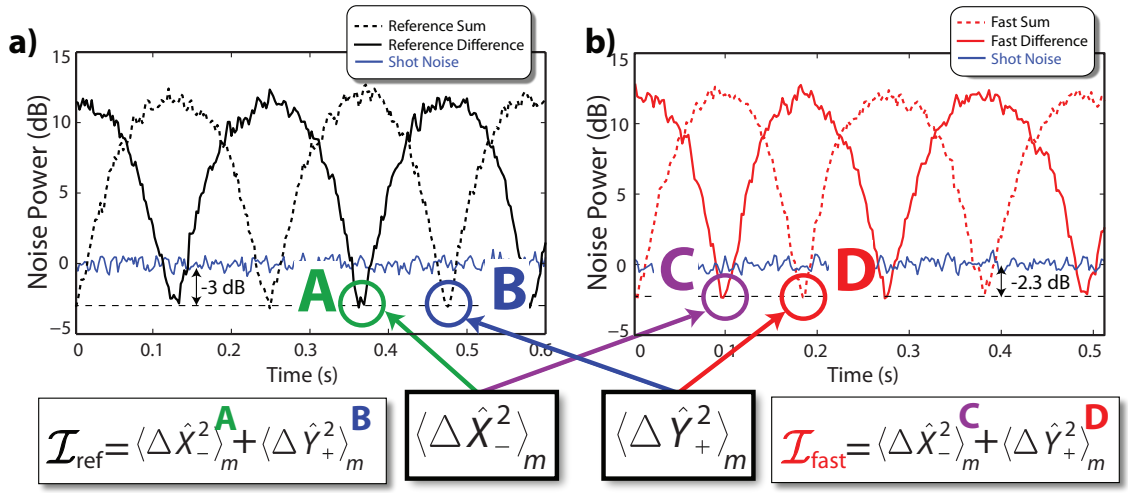


Fig. 4.6: Determination of inseparability $\mathcal{I}$ using the spectrum analyzer (center frequency 750 kHz, resolution bandwidth 30 kHz, sweep time 1 s). Part a) corresponds to the reference case while b) corresponds to data taken with fast light. The minimum noise power of the homodyne difference signal corresponds to squeezing of the $\hat{X}$ quadrature difference (i.e. $\langle \Delta \hat{X}_-^2 \rangle < 1$) while the minimum noise power of the homodyne sum gives the squeezing of phase sum $\langle \Delta \hat{Y}_+^2 \rangle$. We observed the strongest squeezing (lowest value of $\mathcal{I}$) when detecting at 750 kHz. With the fast–light cell inserted into the path of the conjugate, we were able to see $\mathcal{I} < 2$ (setting a resolution bandwidth of 30 kHz throughout) at central detection frequencies ranging from 20 kHz to approximately 3 MHz. The figure shows how the inseparability can be calculated (from the data obtained from the spectrum analyzers) by adding together the values of the appropriate minima and averaging.

each homodyne detector. The spectrum analyzers were configured to measure a central frequency of 750 kHz with a resolution bandwidth of 30 kHz and a sweep time of 10 ms. Although the oscilloscope was also triggered at the same time as the spectrum analyzers, the time traces were only saved upon measuring an average of at least -2 dB of squeezing on either spectrum analyzer. The traces were sorted by whether they corresponded to the local oscillator phases giving rise to squeezing in $\hat{X}_-$ or $\hat{Y}_+$. The oscilloscope traces consisted of 1 million points acquired at a sampling frequency of 2.5 gigasamples per second for a total acquisition time of 400 μs.

We acquired 100 time traces of the homodyne detections of the twin fields for the reference and fast–light cases (yielding a total of 200 traces). The data acquisition time typically lasted a total of about 45 minutes. In between these recordings, we took 5 measurements of shot noise by blocking the vacuum-squeezed twin beams and integrating the power spectral density over the 100 kHz–2 MHz frequency range. No temperature or detuning settings were changed during any given set of measurements. We repeated this procedure over the course of several days to verify reproducibility.

4.4.3 General Comments about the Analysis

The first step in the analysis was to determine which detection (optical sideband) frequencies exhibited quantum correlations indicative of entanglement ($\mathcal{I} < 2$). While scanning the local oscillators, we detected the squeezing of $\langle \Delta \hat{X}_-^2 \rangle$ or $\langle \Delta \hat{Y}_+^2 \rangle$ at different detection frequencies (with a constant resolution bandwidth of

30 kHz throughout). Given any detected squeezing, we calculated the inseparability by summing the normalized noise powers of $\langle \Delta \hat{X}_-^2 \rangle$ and $\langle \Delta \hat{Y}_+^2 \rangle$ on a linear scale to calculate $\mathcal{I}$ (see Fig. 4.6). We observed the strongest squeezing (and therefore smallest values of $\mathcal{I}$) at a central detection frequency of 750 kHz. More generally, we observed $\mathcal{I} < 2$ at detection frequencies ranging from 20 kHz to 3 MHz. Accordingly, we selected a detection frequency of 750 kHz to trigger data acquisitions with the oscilloscope.

4.4.4 Measuring the Cross–Correlation

Since we only observed entanglement at detection frequencies from $\approx$ 20 kHz up to $\approx$ 3 MHz, we were careful to filter out any frequencies where we could not definitively show quantum correlations when computing the temporal cross-correlation function between the beams. In other words, we wanted to prevent any frequencies exhibiting two-mode excess noise from accounting for any possible advance in the cross–correlation function (since we are interested in studying the propagation of quantum information associated with entanglement). Figure 4.7 illustrates the band-pass window that we applied to the probe and conjugate homodyne data prior to computing the cross–correlation function.

4.4.5 Measuring the Inseparability

To study the advance in the squeezing and inseparability, we computed the Fourier transform of the sum or difference of the probe and conjugate homodyne detections (depending on whether the oscilloscope was triggered by detecting squeez-

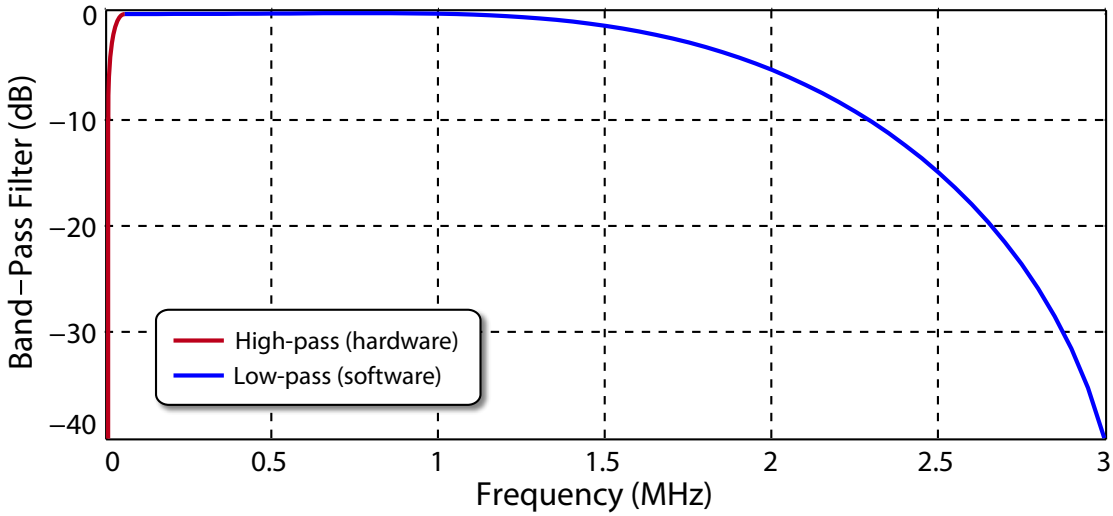


Fig. 4.7: The frequency profile of the effective band-pass filter that was used to filter the probe and conjugate homodyne detections for all three cases (fast–light medium, slow–light medium, and the reference case). The high-pass portion of the filter (at frequencies below 100 kHz, red curve) was achieved in hardware using an LC circuit placed between the output of the detectors and the input of the oscilloscope. The attenuation of the higher frequencies was achieved in software using a Hanning window with a -3 dB cutoff point of 1.75 MHz.

ing of $\langle \Delta \hat{X}_-^2 \rangle$ or $\langle \Delta \hat{Y}_+^2 \rangle$). We then integrated the resulting power spectrum from 100 kHz–2 MHz, well within the bandwidth where we observed entanglement. Selecting a wide bandwidth ensured a large total noise power after the integration, which assisted in repeatably obtaining the same magnitude of advancement. We repeated this process when detecting shot noise in order to normalize the joint noise power obtained with the probe and conjugate. We then shifted the probe and conjugate time traces in time (in steps of 0.4 ns, the sampling period of the oscilloscope) and repeated the analysis over a range of relative delays.

4.4.6 Measuring the Mutual Information

To compute the quantum mutual information, we assume that the initial two–mode squeezed state is Gaussian (i.e. it is completely described by a Gaussian

Wigner function). Furthermore, we assume that the state remains Gaussian after one portion propagates through the second 4WM process (the fast/slow–light cell). In other words, we assume that only Gaussian quantum operations take place in either of these cells. These assumptions simplify the calculations significantly since any Gaussian state is fully characterized by the first and second moments of the field quadratures, which can be obtained from measurements taken with homodyne detectors.

To obtain the excess noises of the individual fields, we Fourier transformed the time traces of the probe and conjugate homodyne detections and integrated the power spectral density over the same 100 kHz–2 MHz bandwidth used to measure the squeezing and inseparability. The two–mode squeezing and the excess noises of the individual beams and are sufficient to compute the state's covariance matrix, which is given by

$$\gamma_{ij} = \frac{1}{2}\langle \hat{R}_i \hat{R}_j + \hat{R}_j \hat{R}_i \rangle - \langle \hat{R}_i \rangle \langle \hat{R}_j \rangle. \tag{4.2}$$

where $R_i \equiv (\hat{X}_i, \hat{Y}_i)$ and $i \in \{p, c\}$ (p denotes the probe and c the conjugate). In this standard form, the on-diagonal sub-matrices characterize the individual modes' fluctuations while the off-diagonal sub-matrices capture the covariances between the two modes' quadrature fluctuations [89]. In a similar fashion to the delay–dependent squeezing, we computed the delay–dependent covariance matrix of our state.

The quantum mutual information, also referred to as von Neumann mutual information, is defined under quantum information theory (see Chapter 3). For a general two–mode state this quantity is defined in terms of the von Neumann

entropy $S_V(\rho) = -\mathrm{Tr}(\rho \log \rho)$ according to

$$I(\rho) = S_V(\rho_1) + S_V(\rho_2) - S_V(\rho). \tag{4.3}$$

Here ρ denotes the full state density matrix and ρ_i denotes the reduced density matrix of the subsystems after the partial trace has been evaluated over the other mode. For the case of a continuous–variable Gaussian state, the calculation of the mutual information involves the symplectic eigenvalues of the standard–form and partially-transposed covariance matrix, as described in more detail elsewhere [86,89]. From the delay–dependent covariance matrix of the state of vacuum–squeezed twin beams, we computed the average delay–dependent quantum mutual information.

The uncertainty in determining the advance of the peak of the mutual information for the fast–light case ($\pm$ 0.1 ns) was determined by computing the standard deviation of the peak advancements over all 200 experiments. To estimate the uncertainty of the leading tail, we computed the standard deviation of the mean of the magnitude of the mutual information at a given delay time and then used standard error propagation techniques to estimate the uncertainty in the timing of the leading tail. This lead to an estimated time uncertainty of no more than 0.7 ns in the leading tail.

4.4.7 Effects of Phase–Insensitive Gain

We provide a rudimentary analysis of how the phase–insensitive gain associated with the fast–light (slow–light) medium would be expected to affect the mutual

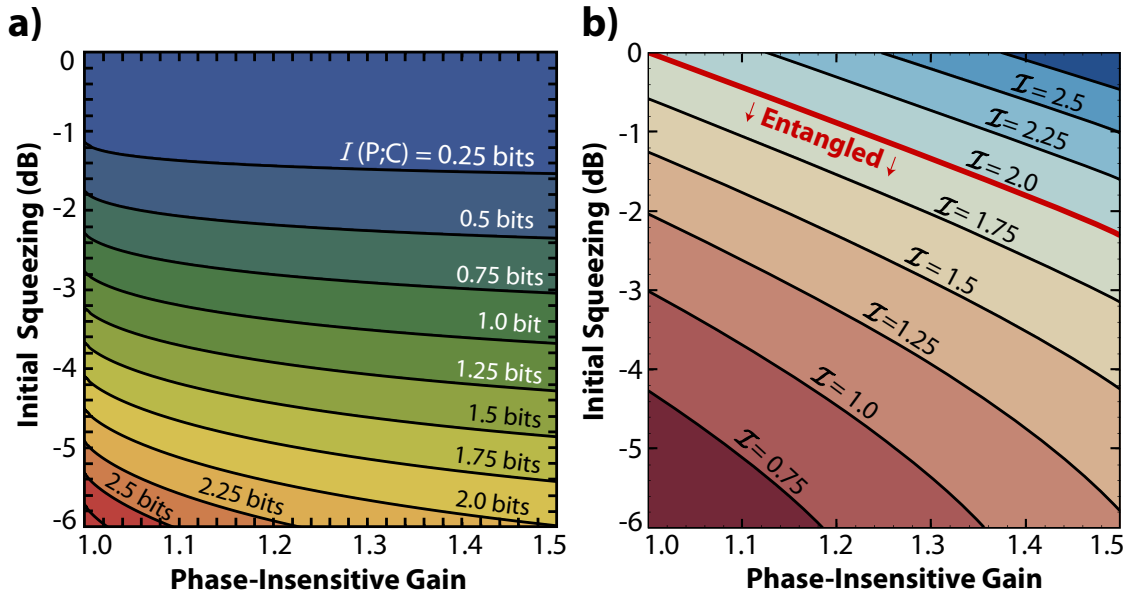


Fig. 4.8: a) Decay of the von Neumann mutual information $I(P;C)$ and b) growth of the inseparability parameter $\mathcal{I}$ when one of the twin beams is subject to a phase-insensitive gain. In this model, the probe and conjugate are assumed to initially be in a pure EPR state. The units in part a) reflect the assumption that the squeezed noise power in the detection bandwidth $\mathcal{B}$ is flat and sampled for a time $\Delta t = \mathcal{B}^{-1}$. The red demarcation in part b) indicates the total phase–insensitive gain required to lose entanglement given some initial two–mode squeezing ($\mathcal{I} < 2 \longleftrightarrow$ entanglement).

information, $I(P;C)$, and inseparability parameter, $\mathcal{I}$. A more complete treatment might incorporate the dispersive properties of the fast–light (slow–light) medium and the squeezed power spectrum to obtain any delay–dependent behavior of the system. Our objective here is only to establish a simple theoretical picture of how the peak mutual information and minimum inseparability would be expected to behave (given certain initial conditions) after one of the entangled beams is sent through a phase–insensitive amplifier.

Assuming that the initial two–mode squeezed vacuum state is ideal (i.e. a pure EPR state), the state's covariance matrix assumes the form

$$\boldsymbol{\gamma} = \begin{pmatrix} \cosh 2r & 0 & \sinh 2r & 0 \\ 0 & \cosh 2r & 0 & -\sinh 2r \\ \sinh 2r & 0 & \cosh 2r & 0 \\ 0 & -\sinh 2r & 0 & \cosh 2r \end{pmatrix}$$

where r is the squeezing parameter and the level of squeezing is given by e^{-2r}. It is well known that passing one mode of this state through a phase–insensitive amplifier will necessarily add noise. The effect of phase-insensitive gain can be described by the transformation [29]

$$\hat{a} \rightarrow \mu\hat{a} + \nu\hat{b}^{\dagger}. \tag{4.4}$$

Here $\hat{a}$ denotes the annihilation operator of one of the twin modes and $\hat{b}^{\dagger}$ the creation operator for the second, unused (i.e. vacuum–seeded) input port of the amplifier. The coefficients are related to the phase-insensitive gain G according to $|\mu|^2 = G$ and $|\nu|^2 = G - 1$. When one of the two EPR modes is subject to an ideal phase–

insensitive amplifier, the covariance matrix of the resulting state can be calculated to be

$$\boldsymbol{\gamma} = \begin{pmatrix} \cosh 2r & 0 & \sqrt{G}\sinh 2r & 0 \\ 0 & \cosh 2r & 0 & -\sqrt{G}\sinh 2r \\ \sqrt{G}\sinh 2r & 0 & 2G\cosh^2 r - 1 & 0 \\ 0 & -\sqrt{G}\sinh 2r & 0 & 2G\cosh^2 r - 1 \end{pmatrix}.$$

The von Neumann mutual information of the twin beams can be calculated from this covariance matrix as a function of the gain G and the initial two–mode squeezing. As illustrated in Fig. 4.8a, phase–insensitive amplification of one of the modes leads to a monotonic decay of the von Neumann mutual information $I(P;C)$ as a function of the gain. A similar procedure can be used to evaluate the amplifier's effect on the inseparability parameter $\mathcal{I}$, which goes as

$$\mathcal{I} = (1+G)\cosh 2r - 2\sqrt{G}\sinh 2r + (G-1). \tag{4.5}$$

Figure 4.8b shows how the introduction of gain leads to an increase in the inseparability parameter $\mathcal{I}$ and, eventually, to a loss of bipartite entanglement ($\mathcal{I} < 2$ is a necessary and sufficient condition to conclude that any Gaussian state is entangled).

Chapter 5: Atomic Ensembles as Quantum Memories

5.1 Introduction

We have seen experimental evidence suggesting that an actively–pumped fast–light medium might not be useful to advance entanglement (or associated quantum mutual information) outside of the light cone[1]. This does not mean that it is not useful to think about using similarly pumped atomic ensembles to delay or store interesting quantum states of light. Storing or delaying quantum states of light presents no obvious challenge to Einstein causality. Moreover, a *quantum memory* would be an essential tool for [101]:

1. Synchronizing resources required for quantum computing (much like Random Access Memory of a classical computer);

2. Converting heralded photons to on–demand photons;

3. Performing entanglement swapping operations vital to practical long–distance entanglement distribution protocols.

[1] Entanglement cannot be used to signal, so it is unclear to me why there would be any *causal* reason that the advancement of entanglement should be forbidden. Still, I would be surprised to see evidence of such an advance in the tails of the mutual information. I am not in possession of any clear physical argument, however, as to what would preclude such a thing from happening.

It is little wonder then that a variety of storage techniques and impressive experimental demonstrations have emerged to tackle the problem [111–116].

In this chapter, we will take up describing some of the essential features of memory techniques involving optically thick atomic ensembles. Although this thesis is really only directly concerned with one particular variant of these techniques (the gradient echo memory (GEM) protocol [117]) ensemble–based memories exhibit striking similarities [118–122]. For example, retrieval efficiencies for these memories are generally limited by the optical depth of the atomic ensemble [122]. Additionally, several of these techniques lend themselves to a convenient "polariton" picture of interacting light and matter waves [123–127], which can be helpful in gaining a deeper conceptual understanding. Finally, the equations of motion describing these techniques are often subject to very similar approximations and limiting behaviors, so comparing them can often be instructive.

5.1.1 Slowly-Varying Envelope Approximation

When propagating in free space (with permittivity ϵ_0 and permeability μ_0), the electric field $\mathbf{E}(\mathbf{r}, t)$ satisfies the wave equation

$$\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{5.1}$$

Since we are assuming narrowband radiation, let us express the field (with carrier frequency ω_0 propagating along $\mathbf{z}$ and polarized along $\hat{\epsilon}$) as

$$\mathbf{E}(\mathbf{r},t) = E(\mathbf{r},t)e^{i(k_0 z-\omega_0 t)}\hat{\epsilon}. \tag{5.2}$$

Upon inserting this expression (suppressing the arguments of $E(\mathbf{r},\, t)$ for notational simplicity) into the wave equation yields

$$\nabla^2 E + 2ik_0\frac{\partial E}{\partial z} - k_0^2 E = \mu_0\epsilon_0\bigg(- \omega_0^2 E + \frac{\partial^2 E}{\partial t^2} - 2i\omega_0\frac{\partial E}{\partial t}\bigg). \tag{5.3}$$

The slowly-varying envelope approximation assumes that the function E($\mathbf{r}$, t) varies much more slowly in time and space than its oscillations at the carrier frequency [46]:

$$\nabla^2 E \ll k_0\frac{\partial E}{\partial z} \tag{5.4}$$

$$\frac{\partial^2 E}{\partial t^2} \ll \omega_0\frac{\partial E}{\partial t}. \tag{5.5}$$

With these approximations (and ignoring diffraction), the wave equation drops to first order in derivatives

$$2ik_0\frac{\partial E}{\partial z} + 2i\omega_0\mu_0\epsilon_0\frac{\partial E}{\partial t} - (k_0^2 - \omega_0^2\mu_0\epsilon_0)E = 0. \tag{5.6}$$

To simplify matters, we will approximate the field as a plane wave and write $E(\mathbf{r},t) = E(z,t)$. Noting that $c^2 = \frac{1}{\mu_0\epsilon_0}$ and invoking the familiar definition of the phase velocity ($\frac{\omega_0}{k_0} = c$), we see that the last term drops out. We are finally left

with the equation

$$\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial z}\right)E(z,t) = 0. \tag{5.7}$$

This same idea can be extended to the propagation of weak fields in a vapor of two–level atoms [128]. We will be interested in the regime where the electric field driving the atomic ensemble is weak enough that we only need consider the linear response of the vapor to the drive. We also expect the induced bulk polarization of the vapor to be along the direction of the applied electric field, and so we express the equations of motion as a scalar differential equation. Finally, we will assume that the medium is not magnetic and has the vacuum permeability $\mu = \mu_0$. Since there are no free charges or currents, Maxwell's equations give

$$\frac{\partial^2}{\partial x^2}E - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}E(z,t) = \frac{\partial^2}{\partial t^2}P(z,t). \tag{5.8}$$

Upon assuming the same conditions as Eqs. 5.4 and 5.5, we find that, to lowest order, this equation reduces to [128, 129]

$$\frac{\partial}{\partial z}E + \frac{1}{c}\frac{\partial}{\partial t}E = i\frac{k_0}{\epsilon_0}P. \tag{5.9}$$

The time Fourier transform of $P(z,t)$ is assumed to satisfy the usual relation

$$P(z,\omega) = \chi(\omega)E(z,\omega). \tag{5.10}$$

We expand $\chi(\omega)$ about ω_0 to obtain

$$\chi(\omega_0+\omega) \approx \chi(\omega_0) + \left[\frac{\partial\chi(\omega)}{\partial\omega}\bigg|_{\omega_0}\right]\omega. \tag{5.11}$$

Noting that

$$\int_{-\infty}^{\infty} \mathrm{d}\omega \frac{\partial\chi(\omega)}{\partial\omega}\bigg|_{\omega_0} \omega E(\omega)e^{-i\omega t} = i\frac{\partial\chi(\omega)}{\partial\omega}\bigg|_{\omega_0}\frac{\partial}{\partial t}\int_{-\infty}^{\infty}\mathrm{d}\omega E(\omega)e^{-i\omega t}, \tag{5.12}$$

we substitute the expansion of Eq. 5.11 into Eq. 5.10 and Fourier transform back to obtain

$$P(z,t) \approx \chi(\omega_0)E(z,t) + \frac{\partial\chi(\omega)}{\partial\omega}\bigg|_{\omega_0}\frac{\partial E(z,t)}{\partial t}. \tag{5.13}$$

The propagation equation can then be approximated as

$$\frac{\partial}{\partial z}E + \frac{1}{v_g}\frac{\partial}{\partial t}E = ik_0P \tag{5.14}$$

where the group velocity v_g takes the form

$$v_g = \frac{c}{1+\omega\frac{\partial\chi(\omega)}{\partial\omega}|_{\omega_0}}. \tag{5.15}$$

Our final goal is to connect the field equation to the coherence term in the atomic density matrix. This proceeds simply by modeling the polarization $P(z,t)$ as the sum of the dipole moments of all N atoms per unit volume V. Using the dipole operator $\hat{\mathcal{D}}$ and assuming a uniform atomic density over the volume, we find

that

$$\begin{aligned} P(z,t) &= \frac{N}{V}\mathrm{Tr}[\hat{\mathcal{D}}\hat{\rho}] \\ &= \frac{N}{V}(\hat{\mathcal{D}}_{eg}\hat{\sigma}_{ge} + \hat{\mathcal{D}}_{ge}\hat{\sigma}_{eg}). \end{aligned} \tag{5.16}$$

Absorbing the carrier wave number and volume into the coupling constant g, we are finally left with the Maxwell–Heisenberg equation[2]

$$\frac{\partial}{\partial z}\hat{E} + \frac{1}{v_g}\frac{\partial}{\partial t}\hat{E} = iNg\hat{\sigma}_{eg}. \tag{5.17}$$

5.1.2 Dicke States

We have developed the mathematical machinery necessary to simply describe continuum quantum fields propagating through atomic media. Now we confront an important conceptual question related to the absorption of light: if a field is absorbed by an ensemble of atoms, in what direction do we expect the field to be emitted when we attempt to recover it? In the case of the cavity field, it is understandable to expect the recovered photons to be emitted back into the cavity mode since it is most strongly coupled to the atoms. If we excite a ground state coherence in an optically thick atomic ensemble without the presence of a cavity, why shouldn't we expect the recovered field to be emitted in any arbitrary direction? This would hardly seem useful as a memory!

One relevant physical insight came in the form of a counter-intuitive prediction

[2] As a side note, this equation has also been referred to as the Maxwell–Bloch equation. To me, the Maxwell–Heisenberg name is a little more intuitive because we are solving for the motion of operators (Heisenberg–Langevin equations) and not the density matrix (optical Bloch equations).

made by Robert Dicke in 1954 [130]. Dicke showed that the behavior of a cloud of excited atoms undergoing spontaneous emission would be expected to change dramatically as the atomic density increased beyond a certain critical point (Fig. 5.1). When the interatomic spacing falls below the center wavelength of the spontaneously emitted photons, it becomes impossible to know which atom is responsible for the emission of any individual photon. Under this condition the coupling of the atomic ensemble to the radiation field is invariant with respect to exchange of any two atoms within the sample [129], leading to a spontaneous phase–locking effect of the atomic dipoles throughout the medium. The result is a short, intense burst of spontaneously emitted photons from the atomic cloud. For a sample of N emitters, the characteristic decay time for the excited state coherence (the natural lifetime, τ) falls to $\frac{\tau}{N}$ while the peak radiative intensity rises by N^2 [129].

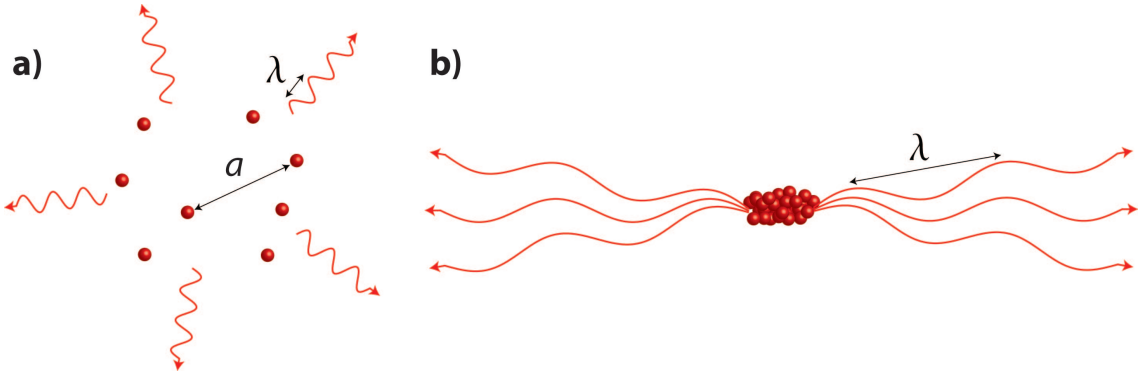


Fig. 5.1: Comparison between a) ordinary spontaneous emission from an ensemble of atoms and b) superradiance. In ordinary spontaneous emission, the average interatomic spacing is a assumed to satisfy $a \gg \lambda$. In contrast, when the atomic cloud is smaller than the emission wavelength, neighboring atomic dipoles undergo spontaneous phase locking, causing the ensemble to radiate faster, more brightly, and anisotropically (depending on the geometry of the sample). For a collection of N atoms, the decay time constant τ decreases to $\frac{\tau}{N}$ while the peak radiated intensity increases by a factor of N^2.

For our purposes, however, the relevant observation made by Dicke is the

anisotropic nature of the emitted radiation. From an initial state free of any dipole correlations, a cooperative ordering emerges akin to "classical antennae radiating in phase throughout the medium" [129]. Although vacuum fluctuations start the superradiative process, the bulk geometry of the radiating system is what determines the asymmetry in the radiation pattern.

In a similar vein, it is possible to establish a preferred direction for the emission of radiation even when the size of the atomic cloud exceeds the radiation wavelength [131]. Moreover, this preferred direction is established by the coherent absorption of the weak incident light field, which is obviously suitable for the purposes of a quantum memory! For example, consider a single photon described by the narrowband wave function $f_k(t)$ incident on an optically thick ensemble of non–degenerate two–level atoms whose bulk size is on the order of the incident mode's center wavelength. Labeling the atomic ground state $|g\rangle$ and the excited state $|e\rangle$, the *collective* state of the ensemble after the photon has been absorbed can be approximated as [101]

$$|\psi\rangle = \sum_{j}^{N} c_j e^{ikz_j} \, |g_1, g_2, ..., e_j, ..., g_N\rangle \, . \tag{5.18}$$

Here k denotes the central wave vector of the incident single photon mode and z_j denotes the position of atom j along the length of the cell. The *entire ensemble* is put into a coherent superposition of singly excited atomic states, and the resulting dipole correlations create conditions reminiscent of superradiance [131]. It is the survival of these dipole correlations written into the ensemble by any single photon

which creates a preferred direction for the spontaneously emitted photon [112]. I will henceforth refer to these important collective states as "Dicke states[3].".

5.1.3 *Three–level Atoms and Spin Waves*

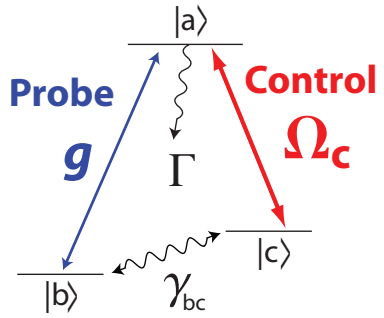


Fig. 5.2: Atomic Λ system coupled by two fields. A "probe field" (coupling constant g) excites the $|b\rangle \rightarrow |a\rangle$ transition while a "control field" (Rabi frequency Ω_c) couples the $|c\rangle \rightarrow |a\rangle$ transition. We will assume a long ground state coherence time compared to the excited state lifetime ($\tau = \gamma_{bc}^{-1} \gg \Gamma^{-1}$).

Let us briefly turn to the behavior of three level atoms in the presence of two light fields (Fig. 5.2). In addition to the light field to be stored (which I will occasionally refer to as the "probe" field), we will consider the role of a strong "control" field whose function is to couple the atoms to the probe. The interaction Hamiltonian governing this system is given by

$$\hat{H} = \underbrace{\omega \hat{a}^\dagger \hat{a} + \omega_{ab}\hat{\sigma}_{aa} + \omega_{bc}\hat{\sigma}_{bb}}_{\text{Bare Hamiltonian}} - \underbrace{ig(\hat{\sigma}_{ab}\hat{a} - \text{H.C.}) + i\Omega_c(t)(\hat{\sigma}_{ac}e^{-i\omega_c t} + \text{H.C.})}_{\text{Interaction}} \quad (5.19)$$

where $\omega_c \equiv 0$. Here ω is the frequency of the cavity mode, g is the familiar cou-

[3] To be a bit more accurate, I suppose that these states should perhaps be referred to as "Dicke–like states" [132] since, unlike the conditions required for superradiance, we do not assume an average atomic separation that is much smaller than the wavelength of emitted radiation [131]

pling constant[4], ω_c is the control field frequency, and Ω_c is the control field's Rabi frequency.

We are now prepared to introduce the concept of a spin wave, a which can be thought of as a collective excitation in the ground state of the atomic ensemble and a traveling quantized light field. For simplicity, we assume that the incident quantum and control fields are perfectly copropagating and thus free of any residual Doppler broadening effects. Introducing the slowly–varying field operators[5] [124] $\hat{\tilde{\sigma}}_{\mu\nu}(z,t)$ and $\hat{\mathcal{E}}(z,t)$

$$\hat{E}^{+}(z,t) = \sqrt{\frac{\omega_{\mu\nu}}{2V}}\hat{\mathcal{E}}(z,t)e^{i\omega_{\mu\nu}(z-ct)/2\pi} \tag{5.20}$$

$$\hat{\sigma}^{j}_{\mu\nu}(t) = \hat{\tilde{\sigma}}^{j}_{\mu\nu}(t)e^{-i\omega_{\mu\nu}(z-ct)/2\pi}, \tag{5.21}$$

we may write down a more convenient expression for the interaction Hamiltonian appropriate for this system. Note that $\omega_{\mu\nu}$ in Eq. 5.20 denotes the resonance frequency of the $|\mu\rangle \leftrightarrow |\nu\rangle$ atomic transition while the superscript j denotes the j^{th} atom in the ensemble. To do so, we again assume (just as with the cavity case) that the atomic density is much greater than the photonic density. Then we can define a quantum field operator for the atoms in a "slice" of the storage volume $A\Delta z$ (where A is the area of the volume in the x–y plane and Δz small enough that $\hat{\sigma}_{\mu\nu}(t)$ does

[4] As a reminder, $g = \mathcal{D}_{ab}\sqrt{\frac{\omega}{V}}\sin kz_0$ for the point $z = z_0$ along the length of the cavity. In this expression, $\mathcal{D}$˙ab is the dipole moment of the atomic transition.

[5] Although I include the tilde when expressing the slowly–varying spin wave $\hat{\tilde{\sigma}}_{\mu\nu}(z,t)$ in this equation, I will suppress tildes in all remaining expressions with the understanding that we are dealing with slowly–varying operators. This way we can avoid tildes and hats over the same operator.

not appreciably change):

$$\hat{\sigma}_{\mu\nu}(z,t) \equiv \frac{1}{N_z}\sum_{j}^{N_z}\hat{\sigma}^{j}_{\mu\nu}(t), \tag{5.22}$$

where we have summed over the N_z atoms in the volume element. We then write the interaction term of the Hamiltonian (Eq. 5.19) by making the replacement $\sum_{j=1}^{N} \rightarrow \frac{N}{L}\int \mathrm{d}z$:

$$\hat{V} = -\int_0^L \frac{\mathrm{d}z}{L}\Big(gN\hat{\sigma}_{ab}(z,t)\hat{\mathcal{E}}(z,t) + \Omega_c(z,t)N\hat{\sigma}_{ac}(z,t) + \mathrm{H.C.}\Big). \tag{5.23}$$

Invoking the Heisenberg-Langevin equations [8] governing the atomic evolution

$$\frac{\partial}{\partial t}\hat{\sigma}_{\mu\nu} = -\gamma_{\mu\nu}\hat{\sigma}_{\mu\nu} + i[\hat{\sigma}_{\mu\nu},\hat{V}] + \hat{F}_{\mu\nu} \tag{5.24}$$

(where $\gamma_{\mu\nu}$ is the decoherence rate and $\hat{F}_{\mu\nu}$ the associated Langevin term) and the field propagation equation (Eq. 5.17), we can solve for the evolution of the atom–field system. There are two Heisenberg equations of motion of particular interest to us:

$$\frac{\partial}{\partial t}\hat{\sigma}_{ba} = -\gamma_{ab}\hat{\sigma}_{ab} + ig\hat{\mathcal{E}}(\hat{\sigma}_{bb} - \hat{\sigma}_{aa}) + i\Omega\hat{\sigma}_{bc} + \hat{F}_{ab} \tag{5.25}$$

$$\frac{\partial}{\partial t}\hat{\sigma}_{bc} = i\Omega^{*}\hat{\sigma}_{ba} - ig\hat{\mathcal{E}}\hat{\sigma}_{ac} + \hat{F}_{bc}. \tag{5.26}$$

To simplify matters, we assume the weak probe limit such that $\hat{\sigma}_{bb} \approx 1$, $\hat{\sigma}_{aa} \approx 0$,

and $\hat{\sigma}_{ac} \ll \hat{\sigma}_{bc}$. Treating Eqs. 5.26 and 5.17 perturbatively in $\hat{\mathcal{E}}$, one obtains [124]

$$\left(\frac{\partial}{\partial t} + v_g \frac{\partial}{\partial z}\right)\hat{\mathcal{E}}(z,t) = \frac{gN}{\Omega^*}\frac{\partial}{\partial t}\hat{\sigma}_{bc}(z,t). \tag{5.27}$$

5.2 Gradient Echo Memory

To adequately describe GEM, we will first take up the task of building an intuition for the physics of photon echoes. From this basic foundation, we will focus the discussion to echo techniques involving controlled inhomogeneous broadening of the absorbing medium through the use of externally applied fields. Finally, we will motivate the polariton picture for GEM which describes the motion of the system's collective normal modes.

5.2.1 Controlled Inhomogeneous Broadening and Photon Echoes

It has been repeatedly argued [121,122] that the ultimate limit for the efficiency of a quantum memory involving an ensemble of atoms is set by the optical depth. Achieving an optical depth near 100% can be accomplished by increasing the density of absorbers and tuning the incident photon closer to resonance. If the photon wave function is exceedingly localized in time, however, the spectral bandwidth of the state might not fit within the natural absorptive linewidth of the atomic medium.

An early proposal [133] to deal with this problem called for taking advantage of the uncontrolled inhomogeneous broadening of an atomic absorption line due to

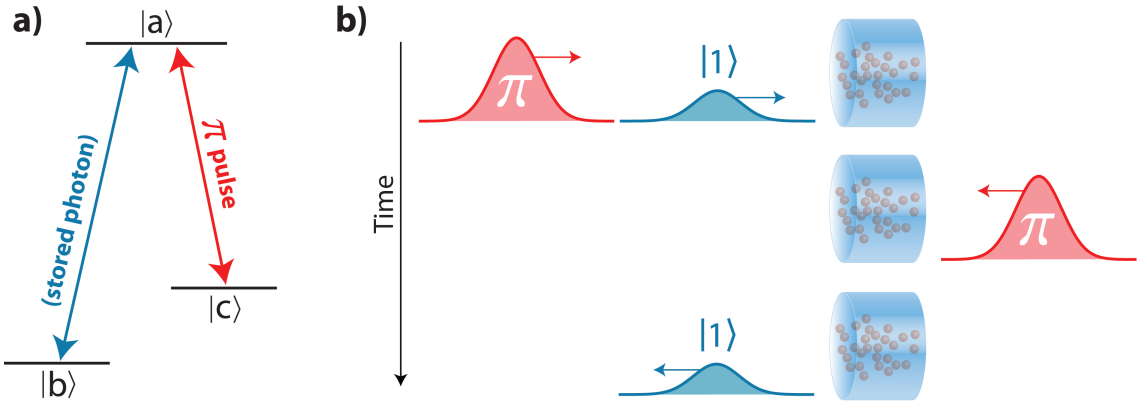


Fig. 5.3: a) Energy level scheme for the earliest proposal to use a photon echo as a quantum memory [133]. All the atoms in the ensemble are assumed to start in the ground state $|b\rangle$, and the ensemble is assumed to be Doppler broadened. The incident single photon is assumed to be resonant with the $|b\rangle \rightarrow |a\rangle$ transition and the bright π pulses are assumed to be resonant with the $|c\rangle \rightarrow |a\rangle$ transition. The $|b\rangle \rightarrow |c\rangle$ transition is assumed not to be dipole allowed. b) Sequence used to store the single photon state. The π pulses used for storage and readout are counterpropagating.

the Doppler effect[6]. The proposal called for a pair of counterpropagating π pulses (Fig. 5.3) which, their claim goes, would lead to a rephasing of the atomic coherence and a photon emission in the backward direction. The π pulse also provided the benefit of moving the excitation to an intermediate metastable state, which would be expected to extend the effective lifetime of the memory.

An alternative approach was later proposed [135] which called for engineering a *controlled*[7] and *reversible* inhomogeneous broadening of the medium's absorption linewidth using external fields. For the remainder of this thesis, we will be interested in such broadening schemes which achieve a spatially dependent center absorption frequency along the length of the cell. One helpful benefit of broadening

[6] The Doppler width $\Delta\omega$ of an absorption line with center frequency ω_0 is given roughly by $\Delta\omega/\omega_0 \approx 1.7\bar{v}/c$ where $\Delta\omega$ is the full width at half maximum of the broadened line [134]. In turn, the average velocity of the atoms $\bar{v}$ is related to the gas temperature T according to $\bar{v} = \sqrt{2k_BT/m}$ where m denotes the atomic mass.

[7] Doppler broadening is inhomogeneous but not "controlled" in the sense that the user does not have any direct control over which classes of atoms absorb certain optical frequencies.

the linewidth in this way is the fact that it allows us to tailor the absorptive properties of the medium in a way which is favorable for the re–emission of the stored field. To see how, consider the Dicke state describing a broadened 2–level storage medium after interaction with the photon [101]

$$|\psi\rangle = \sum_{j}^{N} c_j e^{-i\delta(z_j)\tau} e^{ikz_j} |g_1, g_2, ..., e_j, ..., g_N\rangle . \tag{5.28}$$

Here $\delta(z_j)$ denotes the transition detuning of the j^{th} atom (which is assumed to depend on the atom's position z_j) and τ represents the time after the input field has been absorbed. Although the position–dependent detuning term causes the Dicke state to dephase, simply reversing the sign of the detuning $\delta(z_j) \rightarrow -\delta(z_j)$ implies that the Dicke state should rephase again after time τ [135]. Upon rephasing, the prepared ensemble will collectively act to re–emit the stored pulse in the forward direction. During this emission, the frequency components of the emitted pulse will be detuned from the position–dependent resonance frequency during the pulse's propagation through the remainder of the ensemble. This is obviously ideal since it will help prevent reabsorption of the pulse during its recall.

Let us assume that we apply an external field such that the pulse's detuning from resonance, δ, varies linearly with position (see Fig. 5.4) according to

$$\delta(z,t) = \eta(t)z. \tag{5.29}$$

We have explicitly included a dependence on time in the slope $\eta(t)$ as a reminder

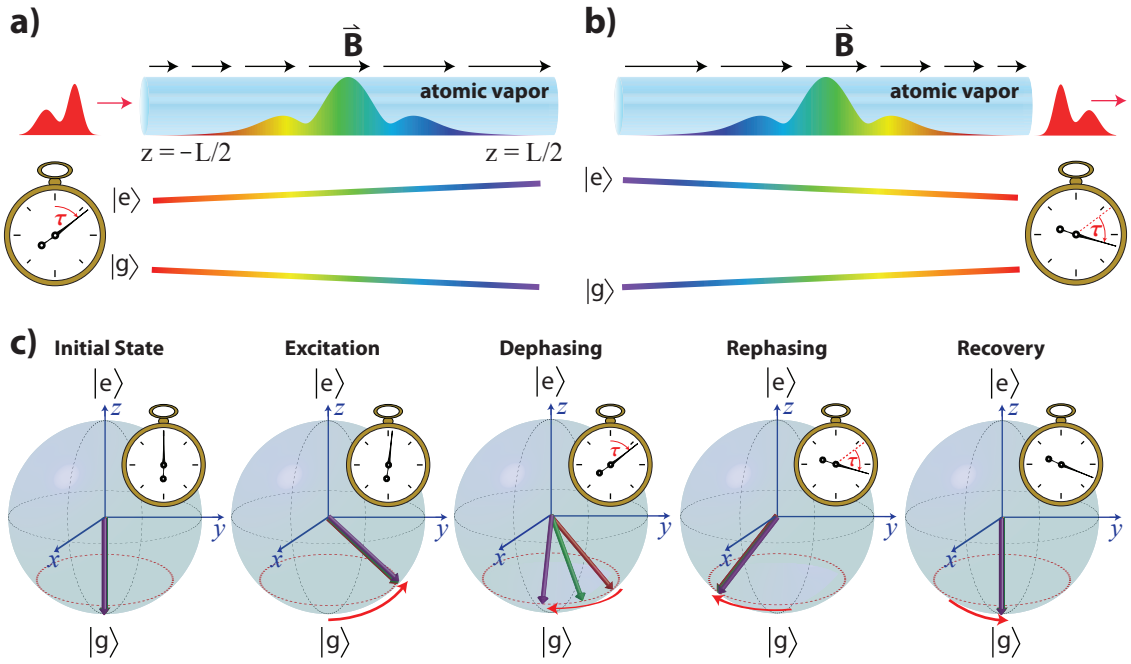


Fig. 5.4: Simple two–level GEM protocol (originally proposed using Stark shifts [117]) as would be implemented using magnetic field gradients. a) The ground and excited states are linearly shifted as a function of position along the length of the cell. This provides the inhomogeneous broadening required to absorb the spectral bandwidth of an incoming pulse. These spectral components are mapped into the atomic coherence along the length of the cell due to spatial dependence of the resonance condition (represented by the rainbow–colored curve of the Fourier transformed input pulse). The resulting atomic spin wave reversibly dephases due to the spatial inhomogeneity of this energy splitting. b) After some time τ (which is assumed to be much shorter than the excited state lifetime Γ^{-1}), the magnetic field gradient is reversed (note that the *direction* of the field is not). This allows the spin wave to rephase, and an "echo" of the stored pulse is recovered. Since earlier input pulses take longer to rephase, the memory is limited to a "first in, last out" order of recall. c) Evolution of the position–dependent Bloch vectors during the process.

that we will be reversing it some time later. The coupled equations[8] governing the evolution of light fields and atomic coherences are given by [117].

$$\frac{\partial}{\partial t}\hat{\sigma}_{ge}(z,t) = -(\Gamma + i\eta(t)z)\hat{\sigma}_{ge}(z,t) + ig\hat{\mathcal{E}}(z,t) + \hat{F}_{ge} \tag{5.30}$$

$$\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial z}\right)\hat{\mathcal{E}}(z,t) = igN\hat{\sigma}_{ge}(z,t) \tag{5.31}$$

where $\hat{\sigma}_{ge}$ denotes the ground-excited coherence. For now, let us make the assumption that the excited state lifetime Γ^{-1} is much longer than the storage and retrieval timescales τ we are studying:

$$\Gamma^{-1} \gg \tau. \tag{5.32}$$

This allows us to ignore the decoherence rate of the excited spin wave and, by extension, the Langevin term $\hat{F}_{ge}$. These equations can be further simplified when expressed in terms of the retarded time[9] $\tau_r = t - \frac{z}{c}$ to yield [117]

$$\frac{\partial}{\partial \tau_r}\hat{\sigma}_{ge}(z,\tau_r) = -i\eta(\tau_r)z\hat{\sigma}_{ge}(z,\tau_r) + ig\hat{\mathcal{E}}(z,\tau_r) \tag{5.33}$$

$$\frac{\partial}{\partial z}\hat{\mathcal{E}}(z,\tau_r) = i\mathcal{N}\hat{\sigma}_{ge}(z,\tau_r) \tag{5.34}$$

where we have defined $\mathcal{N} \equiv \frac{gN}{c}$ to tidy the notation.

[8] These equations make the assumption that $\hat{\sigma}_{ee} - \hat{\sigma}_{gg} \approx 1$ (consistent with a weak quantum field $\hat{\mathcal{E}}(z,t)$ compared to the control field)

[9] The relevant literature [117, 126] describes this change of variables as amounting to transformation into a frame moving at the speed of light. Since such a proposed transformation is not relativistically meaningful, I don't really understand this interpretation. Performing a coordinate transformation to retarded time, however, is perfectly well–defined in any reference frame [39, 48], and this technique has made very similar appearances elsewhere (Eqs. 6.16–6.17 of [129]).

5.2.2 *Polaritons in GEM*

Coupled Eqs. 5.33 and 5.34 can be numerically solved to yield the evolution of the ground state coherence and the electric field as a function of space and retarded time (Fig. 5.5). Although the calculated results are revealing, I personally don't find them to be very intuitive to understand. Fortunately, however, a very clever polariton representation of the atom–field evolution has been developed which helps provide a tidier conceptual depiction of the physics.

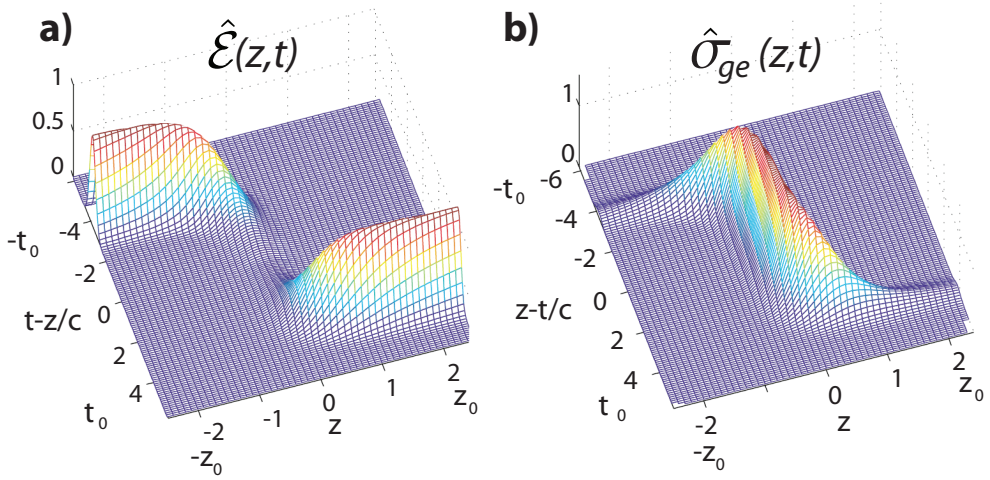


Fig. 5.5: Numerically solved motion for a) $\hat{\mathcal{E}}(z,\tau_r)$ and b) $\hat{\sigma}(z,t)$ performed in [117]. The aspect ratio of the figure has been adjusted to accommodate the format of this thesis. The relevant portion of the figure caption in [117] reads: "The space-time grid plots of the light field intensity and the atomic polarization, respectively. The input pulse duration is $t_{\text{pulse}} = t_0/4$ and the quadrupole induced broadening is $2/t_{\text{pulse}}$."

To find the polariton equations of motion, let us begin by expressing Eqs. 5.33 and 5.34 as a function of k instead of z (i.e. in the spatial Fourier domain). Noting

that the Fourier transform of $xf(x)$ is given by

$$\mathcal{F}\{xf(x)\}(k) = \int_{-\infty}^{\infty} xf(x)e^{-ikx}dx \tag{5.35}$$

$$= i\frac{\partial}{\partial k}\int_{-\infty}^{\infty} f(x)e^{-ikx}dx \tag{5.36}$$

$$= i\frac{\partial}{\partial k}\tilde{f}(k), \tag{5.37}$$

we find that our coupled differential equations take the form (dropping the tildes with the understanding that we are considering Fourier transforms, which are functions of k and not z)

$$\partial_{\tau_r}\hat{\sigma}_{eg}(k,\tau_r) = \eta(\tau_r)\partial_k\hat{\sigma}_{eg}(k,\tau_r) + ig\hat{\mathcal{E}}(k,\tau_r) \tag{5.38}$$

$$k\hat{\mathcal{E}}(k,\tau_r) = \mathcal{N}\hat{\sigma}_{ge}(k,\tau_r). \tag{5.39}$$

With these new equations, we consider the object

$$\hat{\Psi}(k,\tau_r) = k\hat{\mathcal{E}}(k,\tau_r) + \mathcal{N}\hat{\sigma}_{eg}(k,\tau_r) \tag{5.40}$$

subject to the operator $\partial_{\tau_r} - \eta(\tau_r)\partial_k - \frac{ig\mathcal{N}}{k}$. We find that

$$\left(\partial_{\tau_r} - \eta(\tau_r)\partial_k - \frac{ig\mathcal{N}}{k}\right)\hat{\Psi}(k,\tau_r) \tag{5.41}$$

$$= k\partial_{\tau_r}\hat{\mathcal{E}} + N\partial_{\tau_r}\hat{\sigma}_{ge} - \eta\left(\hat{\mathcal{E}} + k\partial_k\hat{\mathcal{E}} + \mathcal{N}\partial_k\hat{\sigma}_{ge}\right) - \frac{ig\mathcal{N}}{k}(k\hat{\mathcal{E}} + \mathcal{N}\hat{\sigma}_{ge}) \tag{5.42}$$

$$= 2\mathcal{N}\partial_{\tau_r}\hat{\sigma}_{ge} - \eta\left(\frac{\mathcal{N}\hat{\sigma}_{ge}}{k} + k\partial_k\left(\frac{\mathcal{N}\hat{\sigma}_{ge}}{k}\right) + \mathcal{N}\partial_k\hat{\sigma}_{ge}\right) - 2ig\mathcal{N}\hat{\mathcal{E}} \tag{5.43}$$

$$= 2\mathcal{N}\left(\partial_{\tau_r}\hat{\sigma}_{ge} - \eta\partial_k\hat{\sigma}_{ge} - ig\hat{\mathcal{E}}\right) = 0. \tag{5.44}$$

We have arrived at an equation of motion governing the evolution of $\hat{\Psi}(k,\tau_r)$ in the Fourier domain. It is useful to observe that the evolution of $\hat{\Psi}(k,\tau_r)$ is governed by the experimentally tunable parameter, $\eta(\tau_r)$. Although previous works have numerically solved for its motion in k–space [125, 126], there is a very simple observation to notice here[10]. Upon inserting the substitution

$$\hat{\Psi}(k,\tau_r) = k^{-ig\mathcal{N}/\eta}\hat{\Theta}(k,\tau_r) \tag{5.45}$$

into Eq. 5.41, we arrive at the revealing equation of motion

$$(\partial_{\tau_r} - \eta\partial_k)\hat{\Theta}(k,\tau_r) = 0. \tag{5.46}$$

This is just the wave equation (consistent with the slowly–varying envelope approximations we have previously made)! We see now that an input light pulse stimulates the soliton–like propagation of the polariton $\hat{\Theta}(k,\tau_r)$ at "velocity" η through k–

[10] Credit where credit is due: I should recognize my friend Justin H. Wilson for noticing the useful change in variables I am about to present.

space [125], which can be understood by inspecting the linearized wave equation in Eq. 5.46. When the gradient is reversed, the velocity η of the polariton's travel through k–space is also reversed and the pulse is eventually recovered.

5.2.3 Λ–GEM

While these ideas are certainly interesting, the thought of implementing them with linear Zeeman shifts in atomic vapors is unrealistic. We have seen, for example, that we are constrained to storage and retrieval timescales which are much shorter than the excited state lifetime (Eq. 5.32). Excited state lifetimes of the alkali metals typically range around tens of nanoseconds [136], so we would need to operate with pulses with temporal widths of a nanosecond at most. For a reasonably–sized vapor cell (say, on the order of 10 cm), achieving Zeeman broadening adequate enough to absorb the pulse's bandwidth would require substantial magnetic field gradient strengths on the order of 100 G cm^{-1}. Not only would this take us out of the regime of linear Zeeman shifts[11], we would be burdened with trying to switch the polarity of such large fields on the nanosecond timescale (which would certainly not be easy).

There is a convenient way to get around these constraints. To do so, let us introduce a third energy level and a strong "control beam" allowing us to couple ground states $|b\rangle$ and $|c\rangle$ (see Fig. 5.6) via a Raman transition. In the limit of a large detunings[12], the photon scattering rate of the driving fields goes approximately as Δ^{-2} [134]. On the other hand, the effective two–photon Rabi frequency of such a

[11] When the Zeeman shift is much smaller than energy splittings due to the hyperfine interaction, F is still a "good" quantum number [84] and the m_F states shift linearly with magnetic field strength and proportional to the m_F quantum number.

[12] Generally speaking, this means $\Delta \gg \Gamma$.

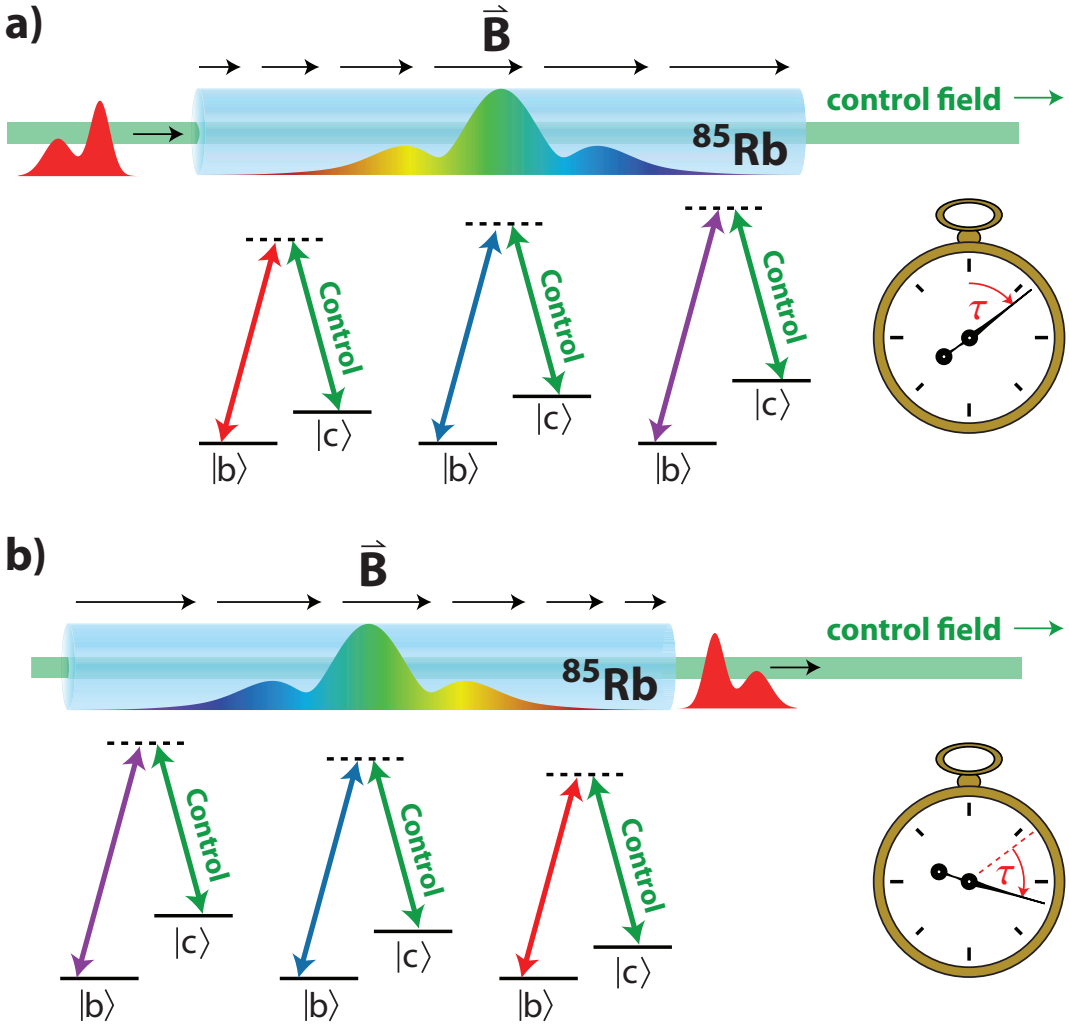


Fig. 5.6: Depiction of the Λ–GEM technique. a) A magnetic field gradient is applied to the atomic ensemble which changes the two–photon resonance condition as a function of propagation distance (the probe pulse and control field are assumed to be copropagating). In this case, level $|b\rangle$ is assumed to be insensitive to magnetic fields while $|c\rangle$ undergoes a linear Zeeman shift. b) When the magnetic field gradient is reversed, the position–dependent two–photon detuning $\delta(z)$ is also reversed and the ground state spin wave rephases.

Raman transition is proportional to Δ^{-1} [134]. Accordingly, if we choose a sufficiently large single photon detuning Δ, we can drive the $|b\rangle \rightarrow |c\rangle$ transition with any desired coupling strength while suppressing absorption to the excited state to some acceptably low level. This last point is crucial: any absorption to the excited state will introduce the deleterious effects of spontaneous emission.

The essential idea behind the Raman scheme is to use Raman coupling of the ground state to the field $\hat{\mathcal{E}}(z,t)$ to store the information contained in the pulse in the long–lived ground state coherence. In addition to allowing for longer storage times, the long–lived ground state coherence provides an important additional benefit. As illustrated in Fig. 5.7, Raman transitions permit absorption linewidths that can be much narrower than the excited state linewidth. Moreover, the Raman transition is used to couple ground states whose energy splitting is nearly Doppler free[13], so the absorption lines are narrow even when working with warm atomic vapors. Accordingly, a Zeeman broadened ensemble of Raman–dressed atoms will more narrowly localize the storage of each constituent frequency of an input pulse. When the echo is retrieved, any re-emitted light is much less likely to be re–absorbed in the remaining path length through the cell since it will be significantly detuned from the medium's absorptive linewidth.

[13] To be free of any Doppler broadening, it is necessary that the ground states being Raman coupled be degenerate (and that the Raman fields be copropagating). For the experiments discussed in this thesis, the ground states are split by the hyperfine interaction (i.e. coupling of the electron's total angular momentum J to the nuclear spin I). The residual Doppler width can be approximated by replacing the ground–excited state transition frequency ω_0 with the frequency difference between the Raman beams $\Delta\omega$. In other words, $\Delta\omega_D \rightarrow 1.7(v_{\rm rms}/c)\Delta\omega$. For a cell temperature of $T \approx 80°C$, we expect residual widths on the order of hundreds of kHz.

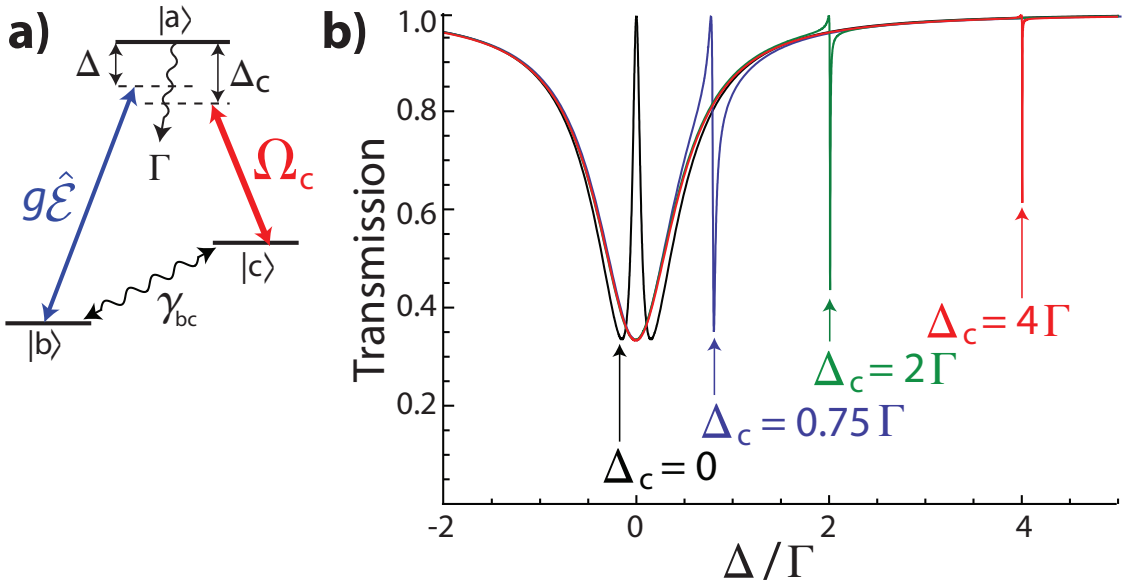


Fig. 5.7: a) Raman scheme used to couple ground states $|b\rangle$ and $|c\rangle$ via the quantum field $\hat{\mathcal{E}}$ and the "classical" control field Ω_c. The ground state coherence time is denoted by $\tau = \gamma_{bc}^{-1}$. b) Theoretical Raman absorption profile (based on [137]) of the weak field $\hat{\mathcal{E}}(z,t)$ propagating through an optically thin medium for different control beam detunings Δ_c. Here I have assumed $\Omega_c = 0.3\Gamma$, $\gamma_{bc} = 0.005\Gamma$, and a propagation length $L = \frac{2\omega_0^2}{3\pi c\Gamma\rho}$ where Γ is the decay rate of the excited state, ρ is the atomic density, and ω_0 is the resonance frequency of the $|c\rangle \rightarrow |a\rangle$ transition. When the control beam is tuned to resonance $\Delta_c = 0$, we satisfy the conditions required to observe EIT when the probe is on resonance ($\Delta = 0$). As the control beam is tuned off of resonance, Raman absorption is observed.

5.2.4 *Adiabatic Elimination*

Does the introduction of a third atomic state and a second optical field (the control field) significantly complicate the mathematical modeling of the physics? Fortunately, these extra degrees of freedom can be effectively removed so that the problem can be mapped back onto the two–level system we have already considered. This should not be surprising since the excited state and control beam are only used as a means to couple the spin wave $\hat{\sigma}_{bc}(z,t)$ to the stored field $\hat{\mathcal{E}}(z,t)$. Ideally, the atoms are never excited to $|a\rangle$ but only pass through the "virtual" intermediate state $|i\rangle$ (Fig. 5.8).

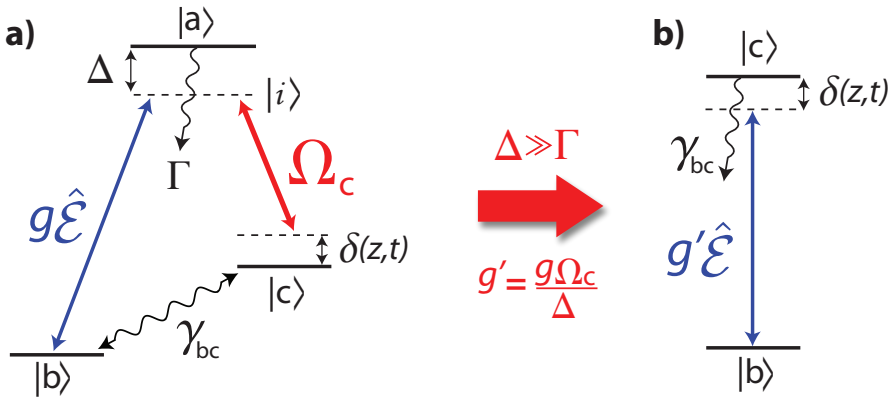


Fig. 5.8: Depiction of adiabatic elimination. Raman transition coupling the ground state assuming a two–photon detuning $\delta(z,t) = \eta(t)z$ that depends on position and time according to the magnetic field gradient. b) When $\Delta \gg \Gamma$, the excited state $|a\rangle$ is effectively removed from the problem. The atomic ensemble and control field act together to mimic an ensemble of two–level atoms with an effective "excited state" lifetime γ_{bc}^{-1} and effective single–photon detuning $\delta(z,t)$. The physical origin of γ_{bc}^{-1} resides in both the single–photon scattering rate as well as atomic diffusion out of the path of the control beam or along the longitudinal axis of the cell. The new coupling strength $g' = \frac{g\Omega_c}{\Delta}$.

Let us make these ideas a bit more precise. Invoking the Heisenberg–Langevin

equation for $\hat{\sigma}_{ac}$ (assuming a real control field Rabi frequency for simplicity)

$$\partial_t \hat{\sigma}_{ac} = -(\gamma + \gamma_{bc} + i\Delta)\hat{\sigma}_{ac} - ig\hat{\mathcal{E}}^\dagger \hat{\sigma}_{bc} + i\Omega_c(\hat{\sigma}_{aa} - \hat{\sigma}_{cc}) + \hat{F}_{ac}, \tag{5.47}$$

we solve for the steady–state solution for $\hat{\sigma}_{ac}$:

$$\hat{\sigma}_{ac} \approx \frac{-ig\hat{\mathcal{E}}^\dagger \hat{\sigma}_{bc}}{\Gamma + i\Delta}. \tag{5.48}$$

To arrive at Eq. 5.48, we have treated $\hat{\mathcal{E}}(z,t)$ perturbatively and made the approximations $\langle \hat{\sigma}_{aa} \rangle \approx \langle \hat{\sigma}_{cc} \rangle \approx 0$. This approximation can be inserted into the Heisenberg–Langevin equation for the ground state coherence (Eq. 5.26 with the inhomogeneous broadening term due to the magnetic field) to yield

$$\partial_t \hat{\sigma}_{bc} = -(\gamma_{bc} + i\eta(t)z)\hat{\sigma}_{bc} + i\Omega_c \hat{\sigma}_{ab} - ig\hat{\mathcal{E}}\hat{\sigma}_{bc} + \hat{F}_{bc}. \tag{5.49}$$

Now we make an additional approximation which simplifies matters significantly: we assume that $\partial_t \hat{\sigma}_{ab} \ll \Delta \hat{\sigma}_{ab}$, which is essentially an assumption that Δ^{-1} is much larger than any timescale for the evolution of the spin wave [138]. Additionally, consistent with the assumption that the Raman process dominates any single–photon absorption, we assume that $\Delta \gg \Gamma$. With these approximations, paired with Eq. 5.27 governing the evolution of the slowly–varying field, it is

straightforward to eliminate the excited state from the problem altogether:

$$\partial_t \hat{\sigma}_{bc} = \left(-\gamma_{bc} + i\eta(t)z - i\frac{\Omega_c^2}{\Delta} \right) \hat{\sigma}_{bc} - i\frac{g\Omega_c}{\Delta}\hat{\mathcal{E}} \tag{5.50}$$

$$(\partial_t + c\partial_z)\hat{\mathcal{E}} = \frac{ic\mathcal{N}}{\Delta}\hat{\mathcal{E}} + i\frac{c\mathcal{N}\Omega_c}{\Delta}\sigma_{bc}. \tag{5.51}$$

The $\frac{\Omega_c^2}{\Delta}$ term can be understood [138] as resulting from the A.C. Stark effect [84], so we will neglect it by shifting to the appropriate rotating frame. Upon performing the variable substitutions $\Omega_c \rightarrow \Omega_c e^{\frac{ig\mathcal{N}t}{\Delta}}$ and $\hat{\mathcal{E}} \rightarrow \hat{\mathcal{E}} e^{ig\mathcal{N}c/\Delta t}$, we eliminate the first term on the right–hand side of Eq. 5.51. Performing a coordinate transformation to retarded time τ_r once again, we arrive at the coupled equations

$$\partial_{\tau_r} \hat{\sigma}_{bc} = (-\gamma_{bc} + i\eta(\tau_r)z)\hat{\sigma}_{bc} - i\frac{g\Omega_c}{\Delta}\hat{\mathcal{E}} \tag{5.52}$$

$$\partial_z \hat{\mathcal{E}} = +i\frac{\mathcal{N}\Omega_c}{\Delta}\sigma_{bc} \tag{5.53}$$

which take the same form as the coupled equations for the two–level GEM protocol (Eqs. 5.33 and 5.34). Thus, in the far–detuned limit $\Delta_c \gg \Gamma$, we expect the physics of the Λ–GEM protocol to be essentially the same as that for a two–level atom with the exception of a new effective coupling strength $g' = \frac{g\Omega_c}{\Delta}$ and an enhanced memory lifetime (due to the long–lived coherence of the ground state compared to that of the excited state).

5.2.5 Λ–GEM in ^{85}Rb Vapor

To be compatible with interesting quantum states of light, the memory must be absorptive at the wavelengths where these states can be generated. We have seen that EPR entangled states can be generated by four–wave mixing in ^{85}Rb vapor (section 6 of Chapter 1), so we certainly find ourselves equipped with a source of non–classical light. Let us briefly consider how a Λ–GEM memory can be implemented using a vapor of ^{85}Rb.

The first task is to isolate a pair of ground states that can be individually addressed by light fields and whose energy splitting can be tuned by a linear Zeeman shift. To achieve this, we apply a bias magnetic field (with a strength ranging between 10–50 G) along the length of a 20 cm vapor cell which lifts the degeneracy of the m_F sublevels of the ground state manifold (Fig. 5.9). The bias field is assumed to be much larger than any magnetic field gradient that will be applied to achieve the requisite inhomogeneous broadening. In the limit of weak magnetic fields[14], the Zeeman shift of the m_F sublevels[15] goes as [136]

$$\Delta E_{\text{Zeeman}} = \mu_B g_F m_F B_z \tag{5.54}$$

where $\mu_B = \frac{e\hbar}{2m_e}$ denotes the Bohr magneton, g_F is the Landé g factor for hyperfine states, and B_z denotes the magnetic field strength along the length of the cell.

[14] By "weak," I mean field strengths that give rise to a Zeeman shift which is much smaller than the energy splitting due to the hyperfine interaction.

[15] Here I will be ignoring the quadratic Zeeman shift, which gives a correction on the order of 10 kHz for the magnetic fields we will be considering.

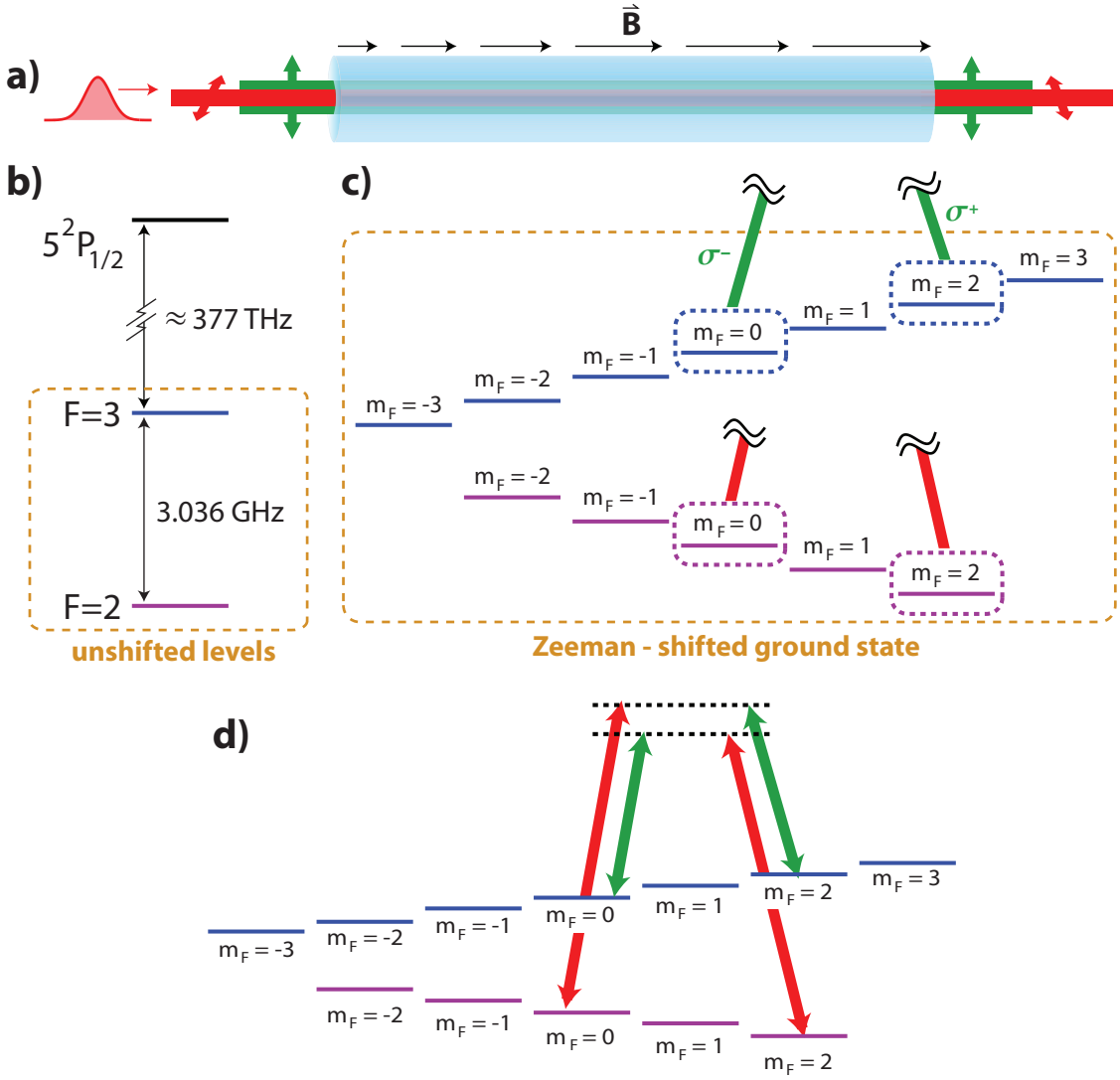


Fig. 5.9: a) Orthogonal linearly polarized Raman beams used in the Λ–GEM experiments presented in this thesis. When propagating along the direction of the magnetic field, a linear polarization can be viewed as a superposition of σ^+ and σ^- polarizations. We therefore simultaneously drive two Raman transitions in any given echo experiment. b) Unshifted hyperfine structure for the D_1 transition in ^{85}Rb. c) A weak bias magnetic field is used to Zeeman shift the hyperfine structure of the ground state. In the weak field limit, the m_F levels are linearly and oppositely shifted in the $F = 2$ and $F = 3$ manifolds. For the Λ–GEM protocol, we coupled $m_F = 2$ and $m_F = 0$ magnetic sublevels. d) Complete depiction of the two Raman transitions.

The hyperfine Landé g factor can be written in terms of the total electron angular momentum J and the nuclear spin I according to [84, 136]

$$g_F = g_J \frac{F(F+1) - I(I+1) + J(J+1)}{2F(F+1)}. \tag{5.55}$$

Since ^{85}Rb has a nuclear spin of $I = \frac{5}{2}$, we find that the g factors for the $F = 2$ and $F = 3$ manifolds are equal and opposite:

$$g_{F=2} = -g_{F=3}. \tag{5.56}$$

This observation is important to notice since equal Zeeman shifts for the sublevels in the $F = 2$ and $F = 3$ manifolds would not allow us to individually address hyperfine transitions of interest.

Due to the equal and opposite Zeeman shifts of the $F = 2$ and $F = 3$ sublevels, there can be more than one resonant Raman transition for oppositely circularly polarized Raman fields. In the case of our experiments, we chose the $m_F = 0$ and $m_F = 2$ ground state pairs for our Raman transitions, giving rise to two coincident Raman resonances illustrated in Fig. 5.9c. Of course, any single Raman transition could in principle be addressed by using copropagating Raman fields with the same circular polarizations. This is an idea we will return to in the next chapter where we will discuss sources of noise that may arise when using cross–polarized Raman beams.

Chapter 6: Storage and Retrieval of Multiplexed Optical Signals in a Gradient Echo Memory

6.1 Introduction

As sketched out in previous chapters, a range of quantum information protocols (e.g. dense coding, teleportation, and certain implementations of quantum key distribution) require entanglement to be shared between users. With this view of entanglement as an information resource, it was perhaps inevitable that a quantum network (or a "quantum internet" [139]) would eventually be proposed, promising a flexible platform for quantum communication [140, 141]. Any attempt to scale such a network over large distances, however, would need to confront the important technical challenges of shuttling quantum information between these nodes or entangling nodes separated by long distances.

Photons are an ideal candidate to carry quantum information between nodes due to their speed of travel and weak mutual interactions[1]. The technical challenge comes when considering the best way to photonically couple these nodes together. Single–mode fibers, a natural choice for an optical conduit, introduce absorption

[1] A qualifier is appropriate here: photons can indeed scatter elastically, even in vacuum [142]. These interactions are rare, however, since they require high energies. What I am referring to are *nonlinear* interactions in air or in optical fibers for typical light intensities of quantum states.

which scales exponentially with propagation distance for sufficiently weak optical powers [134]. In the case of discrete variable (DV) protocols involving single photons as qubits, such absorption leads to two difficulties [143]:

1. Probabilistic loss of the entire qubit (absorption of the photon);

2. Decoherence of the qubit state.

The former problem obviously presents a difficulty since complete absorption of the photon lowers the success probability of connecting neighboring nodes. The second problem is a bit more insidious since even the most forgiving fault–tolerant quantum protocols require minimum gate and state fidelities for proper execution [144–146]. With this in mind, we see that it is not sufficient to repeatedly send single photons down a highly absorptive optical fiber and wait for the occasional detection of a photon on the other end. Absorption damages the photon's quantum information, and it is this information that is ultimately what we are after.

6.2 Quantum Repeaters

To get around these difficulties, Briegel, Dür, Cirac, and Zoller proposed [143] a "quantum repeater" inspired by classically–rooted cascaded error correction techniques. Essentially, quantum repeaters rely on two important operations to be useful: entanglement swapping and entanglement purification [143]. Entanglement swapping can be understood [147, 148] as a special case of quantum teleportation, the process by which an unknown quantum state can be sent from one location to another using a shared entangled resource and classical communication [149]. In

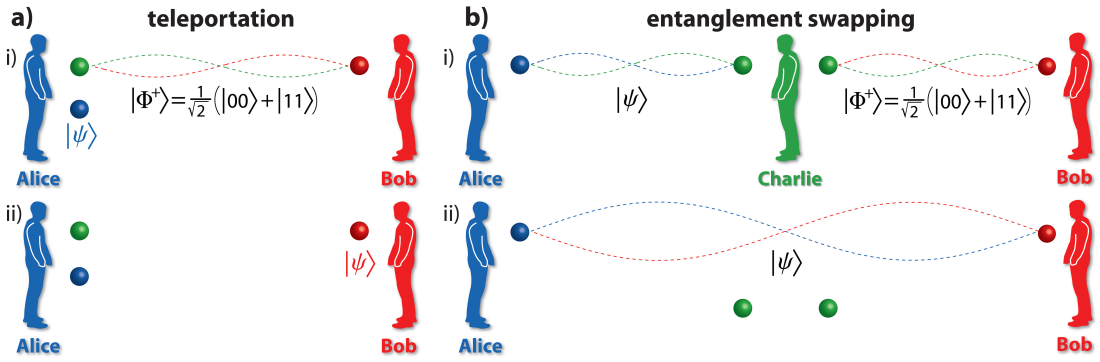


Fig. 6.1: Understanding entanglement swapping as a specific example of teleportation. a) In teleportation, Alice i) sends an unknown state $|\psi\rangle$ to Bob using a shared entangled state (the Bell state $|\Phi^+\rangle$) and classical communication. ii) After teleportation, the state of Alice's particle is transferred to Bob's half of the Bell state. Alice's qubit is no longer left in the state $|\psi\rangle$, so there is no violation of the no–cloning theorem. b) Entanglement swapping can be viewed as i) teleporting Charlie's half of an entangled state $|\psi\rangle$ (which is a bipartite entangled state shared between Alice and Charlie) to Bob. ii) Despite the fact that they started out separable, Alice and Bob's particles become entangled after swapping.

entanglement swapping, rather than simply teleporting the state of a single qubit, half of an entangled state is teleported from one location to another (Fig. 6.1). In this way, it is possible to establish entanglement between systems that never directly interact [150].

Taken alone, however, entanglement swapping is insufficient to provide much of a practical advantage over simple propagation [143]. While the "entanglement fidelity" [14] of *individual* links in the quantum repeater might be high enough to allow for certain fault–tolerant protocols (due to the short distance L separating them), swapping imperfectly entangled states results in further growth of errors. When compounded over many links of the repeater, these errors might eventually yield poor entanglement of the end nodes. To get around this difficulty, a "nested" purification scheme was proposed [143]. If adjacent nodes share a collection of par-

allel bipartite entangled states, these states can be "distilled" down to a smaller set of more highly entangled states [14,151]. Upon swapping entanglement among these distilled states, it should in principle be possible to obtain the required entanglement fidelity between end nodes of the repeater without demanding a burdensome increase in the number of entangled resources[2].

6.2.1 Advantages of a Multiplexed Architecture

The advantages of using multiplexed[3] repeater architecture might already be clear. After all, if a set of parallel bipartite entangled states connecting adjacent nodes must be available for entanglement distillation, then a quantum memory capable of storing these states simultaneously would certainly be convenient. There is another reason, however, that a multiplexed structure would be ideal. The savings in entanglement resources offered by a repeater are offset by a polynomial increase in the time required to entangle repeater nodes [143].

In the same sense that entanglement distillation should allow physicists to get around the entanglement fidelity problem, a multiplexed node architecture can help ease the required time overhead to achieve long–distance entanglement. The essential idea behind a multiplexed approach is to perform many simultaneous attempts

[2] This last point is important. It would be straightforward to obtain many "poorly" entangled states over the entire distance between end nodes using simple lossy propagation without the repeater. Given sufficiently many of such poorly entangled states, it should be possible to distill them down to a much smaller set of entangled states meeting the required entanglement fidelities [14]. The difficulty of this approach, however, is that it would demand an exponential growth in entanglement resources with distance. A repeater, on the other hand, merely demands a polynomial (or even logarithmic) increase in resources with distance [143].

[3] In information networks, multiplexing refers to the common practice of combining multiple signals or data streams into one informationally rich signal propagating over a single information channel.

at entanglement swapping between successive nodes within the connection time $\frac{L}{c}$. The generation of entangled qubits is often a probabilistic process, so performing many swapping attempts increases the likelihood of a success. In particular, it has been shown that temporal multiplexing can relax the time requirements imposed by the probabilistic nature of realistic entanglement swapping [152] and to increase the rate of entanglement distribution [113,153]. Of course, achieving more forgiving time requirements carries the ancillary benefit of lowering the required time–dependent recall fidelity of any realistic quantum memory (and improving the storage time of quantum memories can be a hard problem). Similarly, storage time requirements for quantum repeaters can be reduced by several orders of magnitude if a spatially multiplexed memory is used [154, 155].

6.2.2 Demonstrating Multiplexing with Images

We saw in Chapter 1 that four–wave mixing presents a straightforward method to generate multi-spatial-mode CV entangled states of light near the ^{85}Rb D_1 resonance [98]. In the Fraunhofer limit of diffraction, the entangled probe and conjugate beams exhibit pairwise–correlated coherence areas where quantum correlations are present. These correlated coherence areas can be thought of as providing independently entangled channels which can be processed as separate quantum information resources. Accordingly, these states present a very promising route as a *self–assembled* spatially multiplexed source of distillable entanglement. For this reason, it is worthwhile to investigate the flexibility of a GEM implemented in ^{85}Rb in processing multiplexed signals.

To demonstrate the spatial multiplexing capabilities of GEM, we show that it is possible to selectively retrieve at different times subregions of an image stored in the memory. To enhance the flexibility of this partial retrieval protocol, we also introduce a technique that allows for selective erasure of subregions of the image during storage. By applying an intense beam referred to as an "optical eraser" near the D_1 resonance, we can rapidly induce decoherence in the spin wave through spontaneous emission. First, however, we will turn to the problem of retrieving several images stored in the memory at the same time. These experiments will illustrate the limitations of using a warm vapor as a storage medium and give a sense of whether any significant cross–talk between the image storage might be avoided.

6.3 Temporal Multiplexing

6.3.1 Introduction

In these experiments, the transverse profile of the signal field is shaped using a mask and imaged into the atomic medium. For a control field much wider than the signal, the transverse profile of the collective excitation of the atomic coherence directly mimics the profile of the signal field, so spatial information can be stored in the atomic memory. We first describe the experimental setup and present results on the storage of two images in the atomic medium. We then introduce the criterion of similarity, based on a normalized cross–correlation, to analyze the spatial fidelity of the retrieved images with respect to its input reference image. The crosstalk between

the first and second retrieved images is investigated with regard to this criteria. Finally, we use a resolution chart to measure and quantify the effect of atomic diffusion at a given buffer gas[4] partial pressure on the storage of spatial information. We investigate primarily the transverse spatial diffusion, as longitudinal diffusion along the direction of the field gradient effectively prevents atoms from contributing to the echo. We show that the resolution of the retrieved images primarily depend on the storage duration due to diffusion of the atomic ensemble.

6.3.2 Experimental setup

The experimental setup is shown in Fig. 6.2. The light from a Ti:sapphire laser, blue detuned by 1.5 GHz from the $|5S_{1/2}, F = 2\rangle \rightarrow |5P_{1/2}\rangle$ transition, is used as a control field for the Raman coupling. The signal beam is generated using a double-passed 1.5 GHz acousto-optic modulator to downshift the frequency by the amount of the hyperfine splitting (3.036 GHz). The signal and control beams are overlapped with crossed linear polarizations in the 5 cm long memory cell. The cell contains isotopically pure ^{85}Rb and 667 Pa (5 Torr) of Ne buffer gas and is heated to 80°C. A bias magnetic field of 1 mT (10 G) is applied to the cell to split the ground state Zeeman sublevels and select a specific three-level system. The Raman absorption line is broadened by a 15 μT/cm magnetic field gradient, which we apply in the direction of propagation. The spatial profile of the signal beam is shaped by placing a mask in its path and imaging the mask into the memory cell.

In order to store two different images, two distinct signal beams have to be

[4] The buffer gas is added to decrease the mean free path of the ^{85}Rb atoms, putting the ensemble in the diffusive regime (as opposed to the ballistic regime).

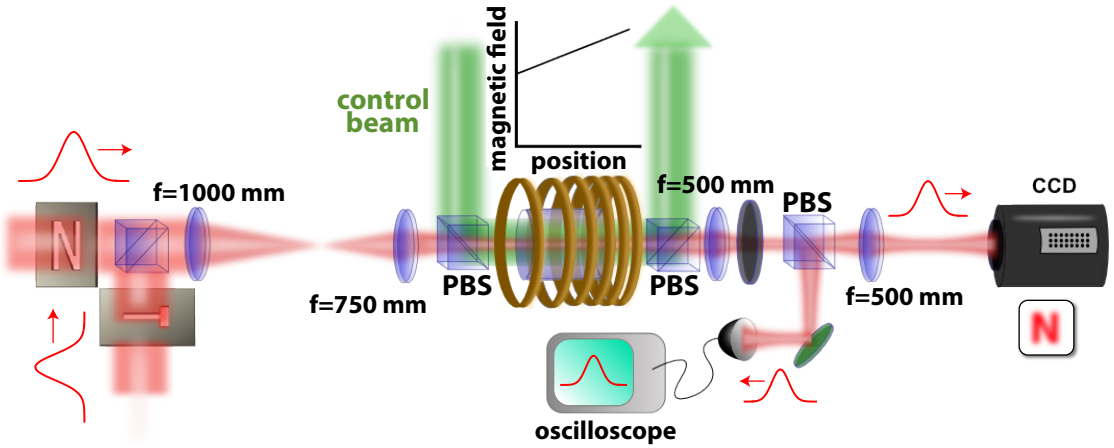


Fig. 6.2: Schematic of the experimental setup for image storage in a GEM. Note that a second gradient coil (wound with the same handedness but with the high–field region at the opposite end of the cell) used to recall the stored images is not pictured. The signal beam is previously shaped temporally with an acousto-optic-modulator (not shown). The spatial profile is fixed using a mask in the beam path and the mask is then imaged into the ^{85}Rb cell using a telescope with a magnification of 0.75 (note that the "N" and "T" masks are placed a focal length away from the first lens). A control field is overlapped with the signal using a polarizing beam splitter (PBS), and the retrieved image is recorded using a fast–gated intensified camera. A small fraction of the signal is sent to a fast photodiode for monitoring the spatially integrated retrieved power. For the storage of two consecutive images, two signal beams are combined with the same polarization on a non-polarizing beam splitter.

shaped independently. The temporal profile of each signal pulse is Gaussian with a full width at $1/e^2$ of 1.1 μs. The storage duration is set by the delay between the input pulse and the flip of the magnetic gradient. The flip duration is on the order of 1.00(1) μs (see Fig. 6.3). After the cell, the control field (which remains on for the entire duration of these experiments) is filtered out with a polarizing beam splitter (with an extinction ratio of approximately 40 dB). A fast–gated intensified camera records the time–integrated light intensity during exposures that last 100 ns. For convenience, a small fraction of the retrieved image beam is sent to a photodiode for recording the retrieved power as a function of time.

6.3.3 Experimental Results

In Fig. 6.3a we present the timeline of the storage and retrieval of two consecutive images. The amplitude of the retrieved pulse is normalized to the input amplitude. In this particular configuration (control beam diameter of 3 mm and power 120 mW, and a probe beam diameter of 1 mm and power 100 μW) we report a retrieval efficiency of 8(2)%. Improvements to this can be obtained by using a longer memory cell, thereby improving the optical depth for the Raman absorption. The optical depth on resonance is approximately 20. The probe beam, however, is far detuned from resonance, so the absorption is due to a 2-photon Raman absorption. In the results presented here, the absorption of the input beam is on the order of 30%, leading to a maximum theoretical retrieval efficiency of 9%[5].

In our experimental procedure, we first checked that a single image can be

[5] This, assumes a recall efficiency that is equal to the storage efficiency and that there is no process leading to gain of the probe field.

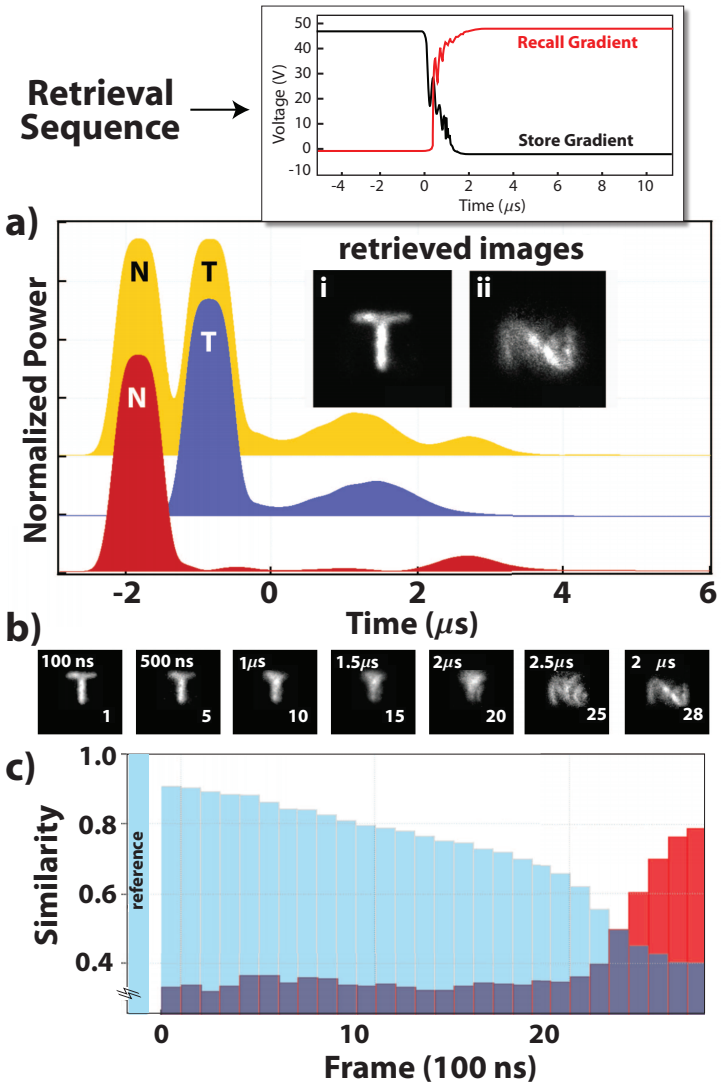


Fig. 6.3: Storage and retrieval of two images in a GEM. a) The spatially integrated intensity from the retrieval of the single letter T (blue) and N (red). The input pulse is at negative time and the retrieved pulse is at the symmetric positive time after the magnetic field gradient flip ($t = 0$). The yellow curve shows the storage of two images, and the curves are vertically displaced for clarity. Retrieved images at time 0.3 μs and 2.7 μs are presented, respectively, in Fig. 6.3a, i and ii. b) The detailed timeline of the retrieved images for 7 frames. The frame number is indicated in the images and the time corresponds to the time after the magnetic flip. (c) S_N (red) and S_T (blue) as function of the frame number. The reference for N or T has a similarity of 1 and is plotted before 0.

stored and optimized the alignment independently for the two different signal beams corresponding to the images of the letters N and T (the red and blue traces in Fig. 6.3a, respectively). To store both images in the atomic memory, the two input pulses are combined on a 50:50 beamsplitter with a delay of 1 μs between them[6]. Two frames from the gated intensified camera of the retrieved images after a storage time[7] of 0.5 μs (frame 3) and 4.5 μs (frame 27) are presented, respectively, in Fig. 6.3a, i and ii. As GEM in this configuration is a first-in-last-out memory [125], the first retrieved image corresponds to the letter T and the second one to the letter N.

A more detailed timeline of 7 frames for the retrieved images is included in Fig. 6.3b. This figure shows the intensity profile of the retrieved light for different times and clearly exhibits an overlap between the two letters in frame 25. To study the overlap quantitatively we define the similarity S to be the cross–correlation of a given frame with a reference image, normalized by the square root of the auto–correlation product :

$$S = \frac{\sum_{i,j} N_{ij}^{in} N_{ij}^{echo}}{\sqrt{\sum_{i,j} (N_{ij}^{in})^2 \sum_{i,j} (N_{ij}^{echo})^2}}, \tag{6.1}$$

where $N_{i,j}^{in}$ is the intensity recorded for pixel $\{i, j\}$ of the input pulse image and $N_{i,j}^{echo}$ for the retrieved image. Defining the similarity in this way (normalized with respect to intensity) was chosen to allow us to confine our attention to the sharpness of the

[6] We were constrained by the short memory lifetime of these experiments to keep the input and recalled pulses as close together in time as possible. Due to the limited absorptive bandwidth of the memory ($\approx$ 1 MHz), this meant confining ourselves to microsecond pulses spaced by the pulse width.

[7] We define the storage time to be the time between input of the peak of a pulse to the time when the peak of the echo emerges from the medium.

retrieved echoes rather than the intensity of the retrieved images. The reference image is selected from among the frames of the input pulse for each letter. In Fig. 6.3c we plot the similarities for the retrieved images with respect to the N and the T reference images (respectively denoted S_N and S_T). For reference, the similarity between the T and the N input images is 35%.

The evolution as a function of time is given for 28 non-overlapping frames representing successive 100 ns time intervals. The first frame is taken 100 ns after the magnetic field gradient is switched at $t = 0$. We temporally expand the retrieved pulse compared to the input (expansion ratio of 1.4) using a different magnetic field gradient for the recovery [125]. This allowed for a slightly longer diffusion time and made it easier to measure its effect in the retrieved images. It is worth repeating that the storage time is given by the time between when peak of the input pulse to be stored enters the memory and when the peak of the retrieved echo emerges.

For the "retrieved" image the initial value of S_T is 88% and it decreases at a rate of 1.1% per frame due to the atomic diffusion that blurs the image. After 19 frames the value of S_T starts to drop faster, at a rate of 4.6% per frame. On the other hand, S_N is initially low (35%) and after 21 frames it suddenly rises at a rate of 8.5% per frame. After frame 28 S_N reaches 78%, indicating the retrieval of a second distinct image. The two curves cross during the 24th frame where $S_T = 47\%$ and $S_N = 51\%$. This is close to a value of 50%, which we take to be a threshold for distinguishing the images.

As shown by the yellow curve of Fig. 6.3a, there is a temporal overlap between the two input images. The similarities corresponding to this overlapping input frame

are $S_T = 55\%$ and $S_N = 48\%$. On the other hand, the frame just before can be identified as an N and the frame just after as a T, with a difference $D = |S_N - S_T| > 0.5$. The retrieved images mimic this behavior with an overlap during frame 24. The value of D immediately before (frame 23) and after (frame 25) is greater than 0.15.

6.3.4 Effect of atomic diffusion

It has been proposed that the decay of S during the storage in an EIT based memory is mainly due to atomic diffusion [156, 157]. We show that our experiment using a GEM technique is consistent with this statement and demonstrate that the expected spatial resolution of the retrieved images can be predicted from the buffer gas partial pressure and the storage time. The internal dynamics of the atoms are described by the optical Bloch equations [8]. In the presence of a buffer gas, the ^{85}Rb atoms are subject to velocity changing collisions with the gas that are assumed to preserve the internal ground state. The atomic excitation $\rho_{bc}(x, y, z, t)$ (which is comprised of both the ground state coherence term in the density matrix and the local atomic density) therefore diffuses in the transverse directions as follows:

$$\frac{\partial}{\partial t}\rho(x, y, z, t) = D\nabla^2\rho(x, y, z, t), \tag{6.2}$$

where D is the diffusion coefficient.

The diffusion along the propagation axis is neglected as it induces only loss in the process and does not affect the normalized contrast of the recovered images.

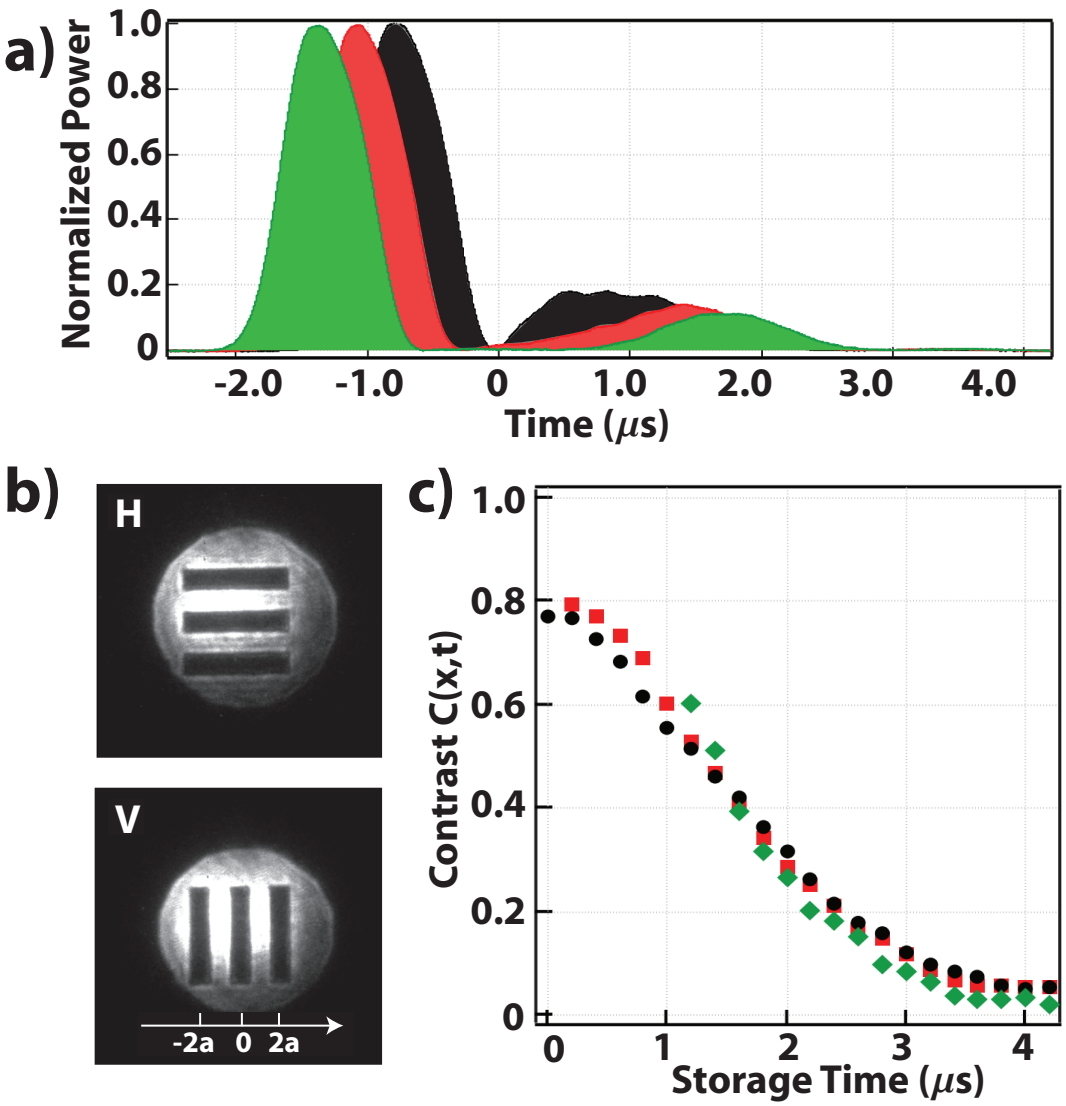


Fig. 6.4: a) The input and retrieved pulses for three different delays. Time t = 0 is taken to be the time of the magnetic field flip. The maximum of the input pulses are at -1.4 μs, -1.1 μs, -0.8 μs (respectively green, red, and black curves). We speculate that the structure on the recalled echo for the shortest storage time (black pulse) is caused by a ringing magnetic field due to inductive flyback for short times after switching the field gradient. b) The input images for horizontal (H) and vertical (V) lines. $1/2a$ is the spatial frequency of the bars, where the distance between the center of the bright and dark lines is a. c) Contrast of a chosen edge between a bright and dark fringe (for the retrieved image of the veritical slit pattern) as function of storage time for 21 frames. The green trace begins at ≈ 1.5 μs since the signal–noise–ratio of the retrieved echo is not large enough for $t < 1.5$ μs. The symbol colors correspond to those in a).

Indeed, if one atom diffuses along the direction of the magnetic gradient, it will not rephase at the same time as others and therefore will not contribute to the expected retrieved image. Using the Green's function formalism, the evolution in time and space of a single point initially at the position $\mathbf{r}' = (x', y')$ can be calculated. Thus the evolution of the atomic excitation from an arbitrary initial condition can be derived by knowing the spatial distribution $\rho(x', y', 0)$ of the excitation at t = 0 as follows:

$$\rho(x, y, t)_{bc} = \frac{1}{\sqrt{4\pi Dt}} \int\int e^{\frac{-(x-x')^2-(y-y')^2}{4Dt}} \rho(x', y', 0)_{bc} \mathrm{d}x' \mathrm{d}y'. \tag{6.3}$$

To evaluate quantitatively the effect of diffusion on the spatial resolution we stored and retrieved images of a resolution test chart consisting of a group of three bars oriented vertically or horizontally as shown in Fig. 6.4b. The resolution chart is imaged into the memory with a magnification of 0.75. For three different pulse delays relative to the magnetic field switching time (Fig. 6.4a) we record 100 ns frames during the retrieval process. The contrast $C(t)$ is defined as:

$$C(t) = \frac{I(x = a, t) - I(x = 0, t)}{I(x = a, t) + I(x = 0, t)}, \tag{6.4}$$

where x can refer to the vertical or the horizontal direction and a is the width of a stored line. At $t = 0$, $x = 0$ refers to the center dark line and $x = a$ to the center of a bright line. $I(x', t)$ is the intensity recorded by the camera at position $x = x'$ (integrated over the extent of the pattern in y) for a frame recorded at time t. With this definition, C will fall below unity for a long storage time as the atoms diffuse

from bright lines to dark lines.

As shown in Fig. 6.4c, the contrast for a given spatial frequency does not depend on the time between the input pulse maximum and the magnetic field switching but on the total storage time. We have also independently checked that the orientation of the lines does not affect the contrast. The fact that the only parameter that affects the contrast is the storage duration is consistent with the hypothesis that diffusion is the only source of the degradation of the resolution.

To further investigate the effect of diffusion, we have compared 4 masks with different spatial frequencies. The resolution of the mask varies from of 1 to 1.6 line pairs/mm, resulting in lines of 375 μm to 240 μm in the image plane in the memory after taking into account the magnification. In Fig. 6.5a, we plot the contrast as a function of storage time for the different masks and the fit of the data to the model of Eq.6.4. In order to take into account independently the quality of our optical system and the contrast degradation due to the memory we have measured the initial contrast of the input pulse C_0. C_0 is measured for the different masks and is a fixed normalization parameter for the fit. We take the value of the diffusion coefficient for 700(20) Pa of Ne buffer gas at 80(5)°C to be 105 cm^2/s as previously reported [158, 159]. The only free parameter we allow for the fits is the initial value of the contrast for different line pair densities, which is limited by the modulation transfer function of our imaging system.

The good agreement between our theoretical model based on atomic diffusion and the experimental data confirms the key role of diffusion in the contrast of the retrieved image for the GEM technique. In Fig. 6.5b we plot the modulation transfer

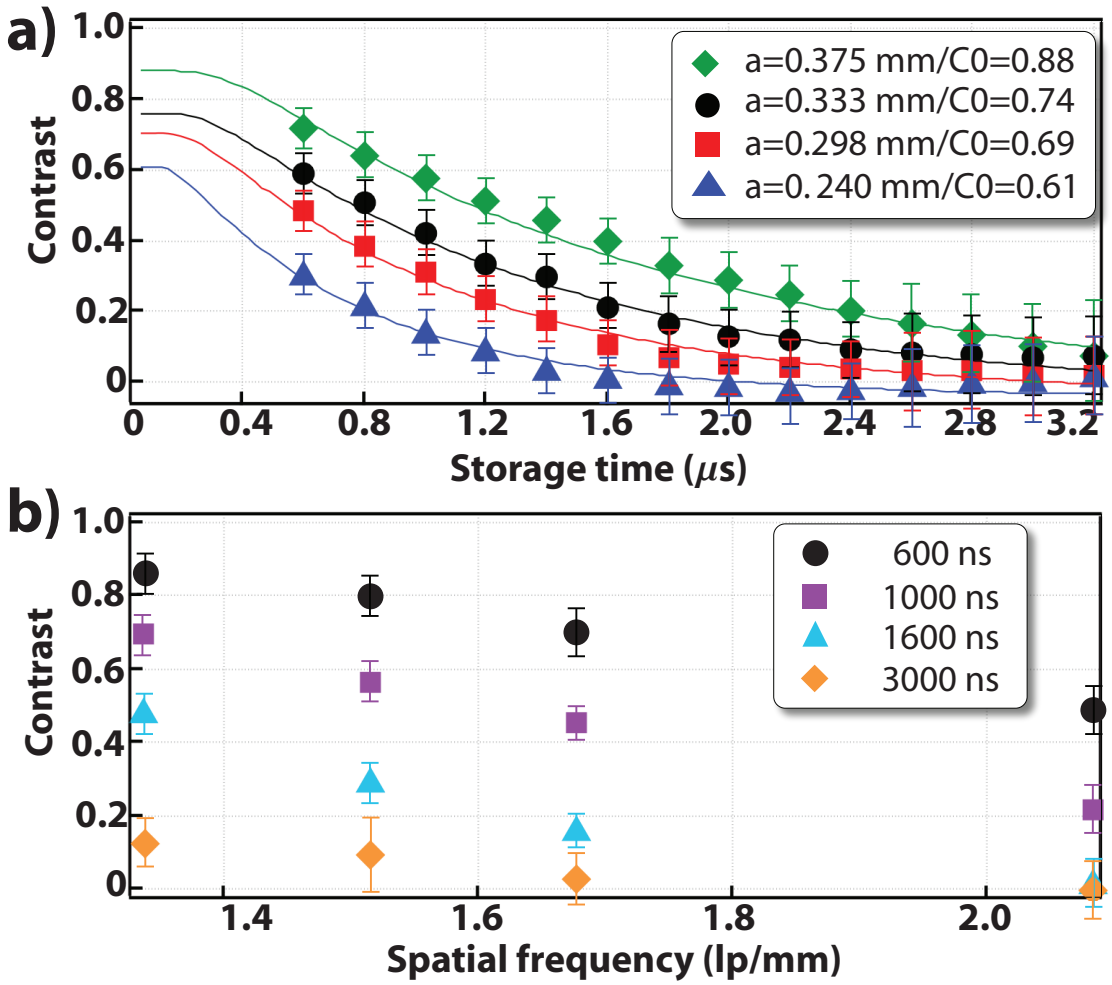


Fig. 6.5: a) Contrast and theoretical fit for four images such as shown in Fig. 6.4b, with bar widths, a, as indicated in the figure. The reduced χ^2 values for the green, black, red, and blue curves are (respectively) computed to be 1.2, .87, .91, and 1.4. The fits are done without free parameters using Eq (3). D = 105 cm^2/s, $t_0 = 0$ and C_0 is given in the legend. b) Modulation transfer function for four different storage times.

function (contrast as a function of spatial frequency) for different storage times. This graph could be a useful benchmark for spatial channel multiplexing for a quantum communication network, as the three lines of the test chart can be seen as three parallel information channels. Fig. 6.5b provides the maximum spatial frequency allowed for a fixed storage time and therefore the number of spatial channels which can be simultaneously used to store quantum information.

6.3.5 Conclusion

We have shown that two images can be stored simultaneously in the ground state spin wave of an atomic ensemble using the GEM protocol. This opens the way to multiplexing simultaneously in time and in space for future quantum memory applications. We confirm that the main limitation of this technique, similar to that in EIT-based memory [156, 157], is the atomic diffusion during the storage time. We expect that this could be overcome by using a cold atomic sample [111, 124] or mitigated by storing the Fourier transform of the image into the memory as suggested in [156, 160]. Finally, as the different spectral components of the input signal are mapped along the length of the cell with GEM, unlike EIT, it would be interesting to investigate if there is an effect of longitudinal diffusion on the signal spectral properties and noise.

6.4 Spatially Targeted Readout and Erasure

We now move on to show that portions of an image written into a gradient echo memory can be individually retrieved or erased on demand, an important step towards processing a spatially multiplexed quantum signal. Targeted retrieval is achieved by locally addressing the transverse plane of the storage medium, a warm ^{85}Rb vapor, with a far-detuned control beam. Spatially addressable erasure is similarly implemented by imaging a bright beam tuned near the ^{85}Rb D_1 line in order to scatter photons and induce decoherence. Under our experimental conditions atomic diffusion is again shown to impose an upper bound on the effective spatial capacity of the memory. The decoherence induced by the optical eraser is characterized and modeled as the response of a two level atom in the presence of a strong driving field.

6.4.1 Experimental set-up

We configure our apparatus to allow three different sub-regions of the atomic excitation to be read out independently (Fig. 6.6). From a master Ti:sapphire laser, a fraction of the light is sent through a fiber-coupled electro-optic phase modulator driven at 3036 MHz corresponding to the hyperfine splitting of the ground state in ^{85}Rb. The first blue optical sideband of the modulated light is selected by injection locking a current-tuned diode laser. The diode light is then pulsed using an 80 MHz acousto-optic modulator (AOM), and the first diffracted order of the AOM is spatially filtered using a single mode fiber. Finally, the mode shape of the light emerging from the fiber is expanded with a telescope and passed through a

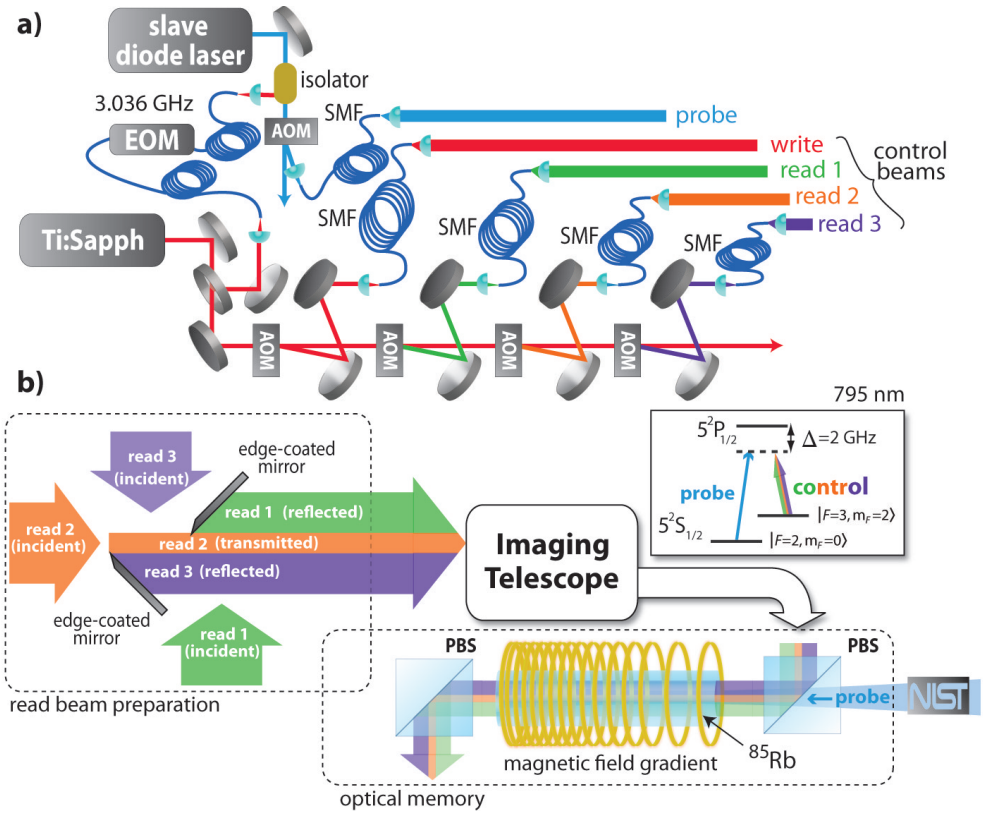


Fig. 6.6: Experimental setup for multiple–readout GEM experiment. a) Generation of the probe, write, and individual read beams. The write control beam is used to write in the image, while the spatially switchable read beam is used to recall different parts of the stored image on demand. The write and read beams are combined on a nonpolarizing beamsplitter (not pictured) immediately after the combination plane of the edge–coated mirrors. A titanium-sapphire (Ti:Sapph) laser serves as a master laser near the control beam frequency. An electro-optic phase modulator (EOM) driven at 3.036 GHz is used to generate optical sidebands in the master laser light, and a slave diode laser is injection locked at the first blue sideband to prepare the probe beam. The probe and control beams are amplitude-modulated using independent acousto-optic modulators (AOMs) and spatially filtered with single mode fibers (SMFs). b) The three read beams are combined using edge-coated mirrors and imaged into the memory cell. The probe and control beams are combined using a polarizing beam splitter (PBS). Inset: Raman coupling of the ground state near the ^{85}Rb D_1 line.

2×8 mm binary intensity mask of the NIST logo. The mask is imaged into the atomic memory with a magnification of 1.25[8]. We refer to this blue-detuned image as the probe.

The remaining master laser power is sent through a chain of four additional AOMs all driven at 80 MHz, and the first diffracted order from each AOM is coupled into separate single mode fibers (Fig. 6.6a). Independently driving these AOMs allows us to implement an arbitrary pulse sequence among the beams while maintaining control of their optical frequencies (which we set to be the same). We use one of these control beams, referred to as the "write" beam, to mediate the Raman coupling of the probe pulse to the atomic ground state coherence. This 50 mW write beam is spatially filtered using a single mode fiber, collimated to a $\frac{1}{e^2}$ diameter of 2 cm, and combined with the probe beam on a polarizing beamsplitter as was done in [161].

The three remaining "read" control beams address different subregions of the transverse plane of the vapor cell at different times. These beams are prepared at the same optical frequency to ensure that all retrieved pulses emerge at the same frequency. This approach simplifies any desired homodyne detection in future experiments by requiring only a single local oscillator frequency to detect the field quadratures of each retrieved subregion. To achieve time-dependent spatial addressability, the optical paths of the beams are combined using an assembly of edge-coated mirrors ("D mirrors") as illustrated in Fig. 6.6b. The combination plane is imaged into

[8] We increased the magnification from 0.75 in the previous experiment where we stored images of the letters "N" and "T" so that we could more finely address spatial features of the image for targeted recall.

the cell in order to obtain two sharply defined boundaries between the three zones of the image that are to be independently retrieved at different times. Note that while the combination plane is not perpendicular to the propagation direction, the mismatch is unimportant since we use a telescope with a final focal length of 1 m, providing a very large depth of field for the imaging system.

In all experiments, we Zeeman-broaden the Raman line to a width of ≈ 1 MHz. All probe pulses were temporally shaped as Gaussian functions with a $\frac{1}{e^2}$ width of 2 μs. The write beam is switched off during the 2 μs of storage time, and portions of the image are sequentially retrieved using three read beams each pulsed for 500 ns (including rise/fall times of 100 ns) centered around the rephasing time of the spin wave (Fig. 6.7a). The atomic memory itself consisted of a 20 cm long ^{85}Rb cell with 1.33 kPa (10 Torr) of Kr buffer gas when heated to approximately 80°C. Similar to the procedure of [162, 163], we choose a relatively high partial pressure of the Kr buffer gas in order to mitigate the blurring effect observed on stored images due to atomic diffusion [161, 164]. Using a technique identical to [161], we estimate a diffusion coefficient for these experiments to be 35 cm^2s^{-1}. The echoes are captured using a gated intensified CCD camera in order to study the time dynamics of the readout.

6.4.2 Results: Targeted Readout

The longitudinal orientation of the spin wave allows any transverse intensity modulation of the stored pulse to be recovered in the echo [161]. Similarly, we are free to modulate the transverse intensity profile of the read beam in order to selectively

read out any desired portion of a stored image. Three subregions of the NIST logo retrieved by the spatially addressable read beams are portrayed in Fig. 6.7b. These subregions were retrieved piecewise during a single ≈ 2 μs rephasing process and not from rephasing the spin wave three times [165] or performing three separate echo experiments.

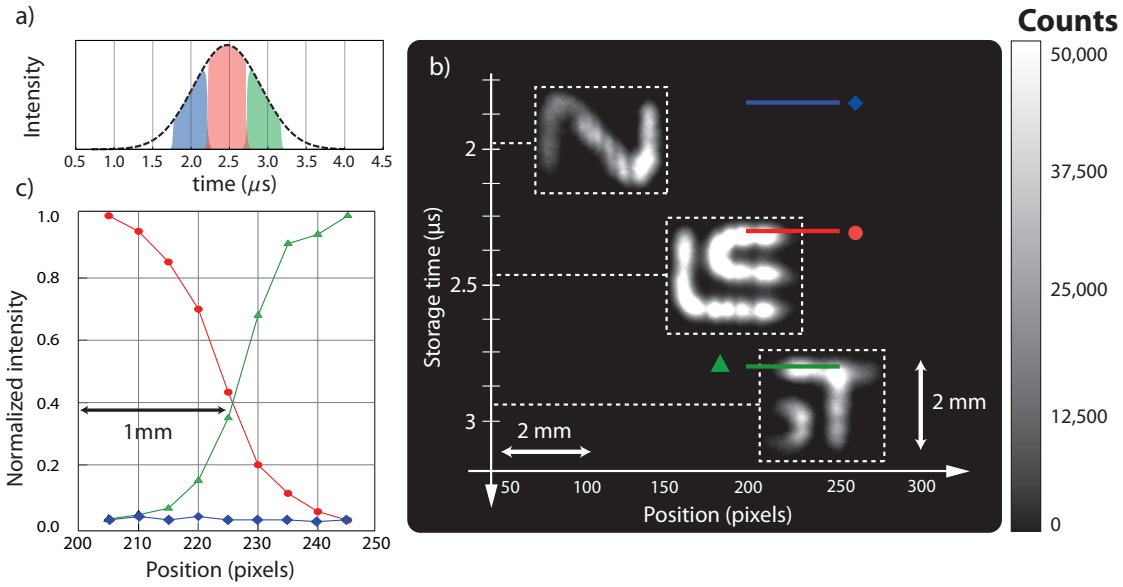


Fig. 6.7: a) Retrieval sequence as a function of storage time. Regions shaded blue, red, and green respectively denote the first, second, and third recalled subregions of the stored image. The black dotted line indicates the typical temporal shape of a single retrieved echo. b) Retrieved portions of the NIST logo as function of storage time and horizontal position. The NIST logo was stored (see Fig. 6.9) once as a single input using a single pulse of the write beam. The image was then retrieved piecewise at various times (2, 2.5 and 3 μs) using three different read beams separately activated during the approximately 3 μs of rephasing time of the spin wave. c) Profile plot of the normalized intensity for a selected subset of pixels (pixels 205–245) in order to show the boundary between two read-out regions: blue diamonds denote the retrieval at 2 μs, red circles at 2.5 μs, and green triangles at 3 μs.

It is apparent from Fig. 6.7a that the second retrieved subregion of the image is brighter than the first and last retrievals. This effect is due to nonuniform illumination of the mask with the probe beam and to the temporal shape of the

retrieved pulse. Although the probe beam was expanded and collimated before passing through the intensity mask, its Gaussian profile contributed to a 35±3% reduction in intensity near the edges of the mask. Since the temporal shape of the echo is also Gaussian with a maximum intensity corresponding to a retrieval time of 2.5 μs, the first and last retrievals exhibited up to an additional 20±3% reduction in intensity.

To characterize the addressability of the memory associated with this technique, we include plots of the normalized intensity profile for a single row of pixels at three different readout times in Fig. 6.7b. The profile plots illustrate the extent to which a continuous intensity profile (the connection between the letter "S" and the letter "T") written into the memory can be sharply separated between two readouts. The normalized intensity profiles recalled in two neighboring readouts show a decline from 0.9 to 0.1 over a length of ≈ 900 μm.

Our strategy of applying several localized read beams during a single rephasing process introduces an effective loss for each individually–recalled subregion of the stored image. In the context of storing and retrieving a spatially multiplexed quantum signal such loss can be avoided by repeatedly inverting the magnetic field gradient as shown in [165] and applying the localized read beams during the entirety of each rephasing time. This approach would extend the total recall time for the experiment, rendering the blurring effects due to the diffusive motion of the atoms to be more significant during each successive targeted retrieval [161, 166].

Taking the diffusive motion of the atoms into account [167], we estimate the number of channels N_{max} of a spatially–multiplexed signal that can be stored and

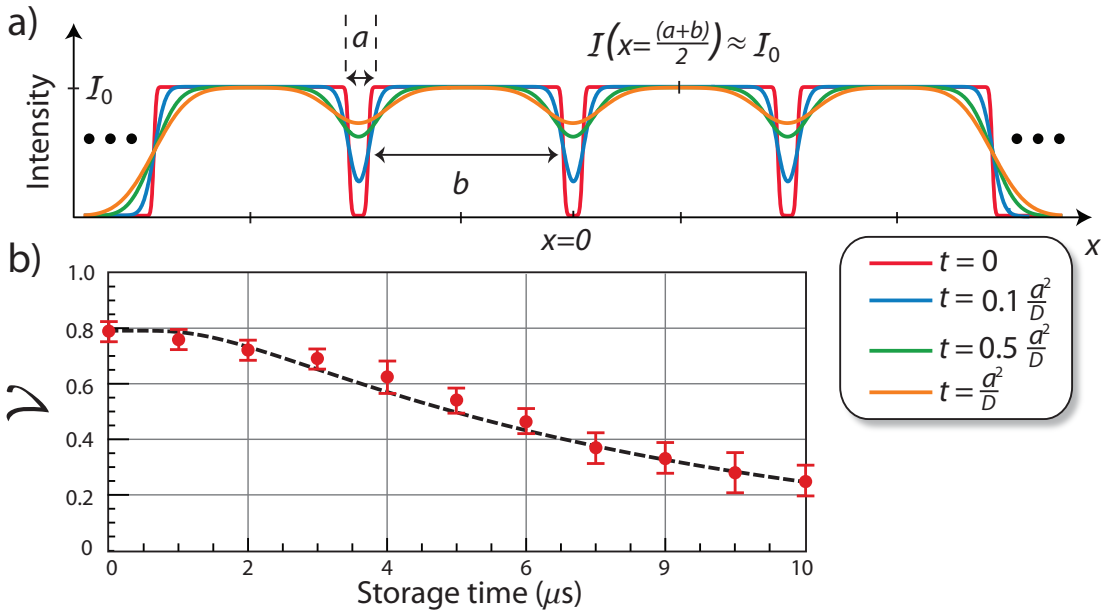


Fig. 6.8: (a) Effect of diffusion on independent channels of width b, separation a, and diffusion coefficient D at various storage times t assuming no background. (b) Observed decay of the contrast (red points) for a stored resolution target at 1 line pair per millimeter. The data is compared against an numerical simulation of diffusion assuming a diffusion coefficient $D \approx 35$ cm^{2}s^{-1}, equal bright and dark stripe widths ($b = a$), and a normalized background intensity of 0.1. The reduced χ^2 of the model is evaluated to be 1.1, corresponding to a P value of 0.36.

independently read–out over time. For simplicity, we consider the storage of independent channels of width b buffered by channel separations of width a (Fig. 6.8). Defining the origin of our coordinate system at the center of the channel separation, we quantify the expected channel independence under the influence of atomic diffusion according to the contrast (Eq. 6.4).

By defining a threshold on the minimum acceptable contrast C_{lim}, we may estimate the maximum allowable linear channel density Λ. Given our empirically determined diffusion coefficient $D \approx 35\ \mathrm{cm^2 s^{-1}}$ and for a $C_{\mathrm{lim}} = 0.9$, we estimate that our memory can accommodate a linear channel density of up to $\Lambda \approx 7\ \mathrm{cm^{-1}}$ for up to three readouts assuming a total storage time of 15 μs (t=15 μs in Eq. 6.3). This density falls well within the resolution afforded by our read beams. We conclude that, for μs storage times, the upper bound on Λ is limited by atomic diffusion.

6.4.3 Results: Decoherence–Induced Erasure

The evolution of the probe optical field $\mathcal{E}$ and atomic ground state coherence $\sigma_{2,3}$ can be described by the coupled Maxwell-Heisenberg equations [168]:

$$\begin{aligned} \partial_t \sigma_{2,3}(z,t) &= -\left(\gamma_{2,3} + i\eta z\right)\sigma_{2,3}(z,t) + ig\mathcal{E}(z,t), \\ \partial_z \mathcal{E}(z,t) &= \frac{i\mathcal{N}\Omega}{\Delta_w}\sigma_{2,3}(z,t), \end{aligned} \tag{6.5}$$

where $\mathcal{E}(z,t)$ is the slowly-varying electric field of the probe, $\sigma_{2,3}(z,t)$ the ground state coherence[9], $\mathcal{N}$ the linear atomic density, g the Raman coupling strength, Ω the

[9] The subscript denotes the fact that we are driving transitions between the $F = 2$ and $f = 3$ ground state manifolds as explained in the previous chapter.

Rabi frequency of the control beam, and Δ_w the detuning of the control beam from resonance. If the ground state decoherence rate $\gamma_{2,3}$ is assumed small compared to the time scale of the experiment, the general behavior of the system is to coherently transfer an optical excitation into a ground state spin wave and back [168]. Accordingly, any decoherence process that damages the spin wave during storage will directly result in loss of the retrieved probe pulse.

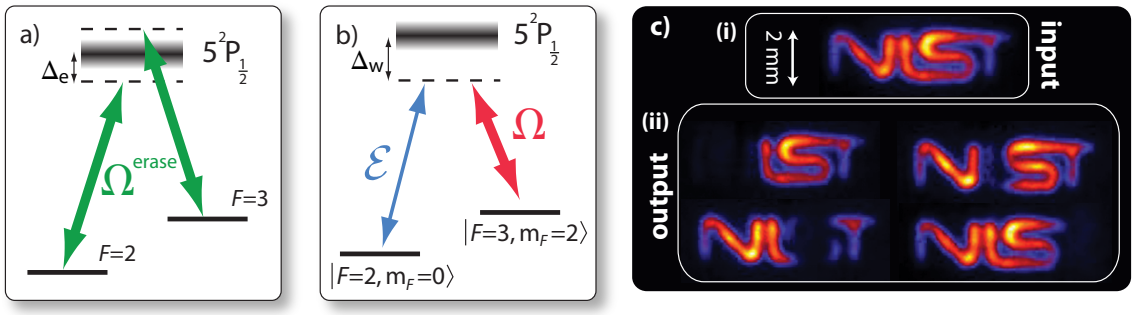


Fig. 6.9: Local erasure of an image. a) Optical eraser applied between the hyperfine ground state splitting with detuning $\Delta_e \approx 1.5$ GHz. (b) Raman absorption scheme used to excite the spin wave. Write beam detuning $\Delta_w \approx 2$ GHz. (c) (i) Retrieved image after 3 μs of storage without erasure. (ii) Retrieved images after 3 μs storage times and local erasure of individual letters of the NIST logo.

We intentionally induce decoherence of the ground state spin wave by pulsing a 10 mW "optical eraser" detuned by $\Delta_e \approx 1.5$ GHz from the D_1 resonance during storage (Fig. 6.9a). Choosing equal detunings for both ground states is meant to minimize optical pumping effects in the Λ-system, allowing us to approximate the storage medium's response as that of a two-level atom. We therefore expect the characteristic decoherence rate $\gamma_{2,3}$ to scale with the scattering rate:

$$\mathcal{R}_{sc} \approx \frac{\Gamma}{2}\left(\frac{\mathcal{I}}{\mathcal{I}_{sat}}\right)\left(1+4\left(\frac{\Delta_e}{\Gamma}\right)^2+\frac{\mathcal{I}}{\mathcal{I}_{sat}}\right)^{-1}, \tag{6.6}$$

where $\Gamma = 2\pi \cdot 5.75$ MHz is the $5^2P_{\frac{1}{2}}$ excited state linewidth, $\mathcal{I}_{sat}$ is the resonant saturation intensity of the D_1 line[10], and $\mathcal{I}$ is the optical eraser beam intensity. The decoherence of the ground state $\gamma_{2,3}$ therefore depends on the detuning, optical power, and integrated intensity over the pulse duration of optical eraser.

In the presented experiments (where we demonstrate the localized erasure capabilities of our memory), the write beam is also used as a read beam (Fig. 6.9b) in order to attempt the recall of the entire stored image in a given experiment and to ensure identical conditions during storage and retrieval. We prepare the optical eraser by replacing the optical fiber previously used for the central read beam ("read 2" in Fig. 6.6b) with an independent diode laser tuned to $\Delta_e \approx 1.5$ GHz. Fig. 6.9c demonstrates how letters of the NIST logo can be individually erased by translating the edge–coated mirrors used to generate the spatially–dependent read beam.

To quantitatively study effects of the eraser, we store a resolution target pattern with a density of 1.5 line pairs per mm (see Fig. 6.10a). Since the edge-coated mirror assembly was mounted on a translation stage, our setup facilitated easily controllable targeting of 2x3 mm regions of interest in the fringe pattern. In particular, we found that erasure could be strongly localized to a particular region of interest, allowing for the complete removal of a selected stripe with negligible effects on neighboring stripes (Fig. 6.10b). This degree of precision is consistent with the resolution limits achieved in this experiment, suggesting that this erasure technique could be used to delete undesired channels of a densely spatially–multiplexed signal.

To study the time scale necessary for complete erasure, we measured the re-

[10] For ^{85}Rb, $\mathcal{I}_{\rm sat} = 4.4876(31)$ for π polarized light resonant with the D_1 line [136].

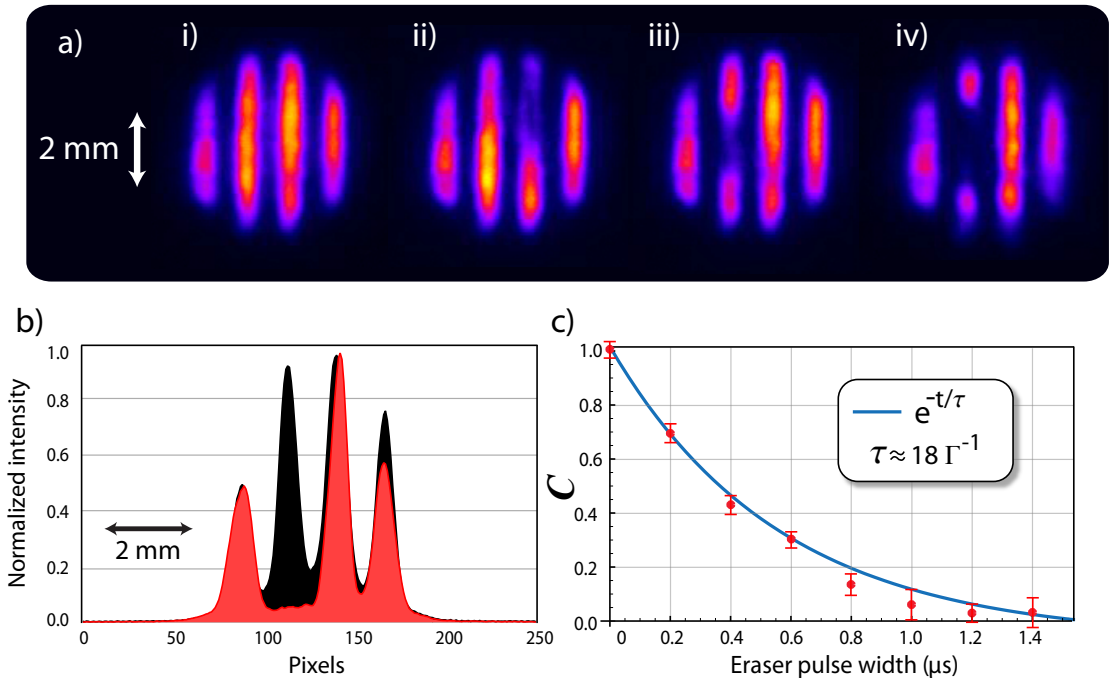


Fig. 6.10: Localized erasure of a stored image. a) Retrieved patterns (1.5 line pairs per mm) for the case of (i) no optical eraser, (ii) and (iii) a 600 ns eraser pulse, and (iv) a 1 μs eraser pulse. b) Profile plot of the retrieved intensity without erasure (black) and after exposure to a 1 μs eraser pulse (red). c) Contrast C of the retrieved fringe patterns (red points) versus pulse width of the optical eraser. The solid blue curve is the modeled decay of C predicted by the scattering rate of a two-level atom for a detuning of Δ=1.5 GHz, which yields a reduced χ^2 value of 1.43 (the best fit decay time constant is 17 Γ^{-1}). In all cases the storage time was fixed to 3 μs.

trieved contrast C subject to a variety of eraser pulse widths. All fringe patterns were stored for 3 μs to ensure that any effects due to diffusion were common among all shots. When evaluating C in a given shot, the fringe patterns were normalized to the brightest undeleted fringe in order to correct for shot-to-shot laser fluctuations. Additionally, an average background from the read beam was subtracted from each shot to allow C to approach zero in the limit of large eraser pulse widths. The measured decay of C, as well as the expected decay for a two-level atom, are plotted against storage time in Fig. 6.10c. Using the model of a two-level atom, we estimate a decay rate of $\approx 18(1)\ \Gamma^{-1}$. This very simple model shows close agreement with a best-fit exponential decay rate of $\approx 17\ \Gamma^{-1}$.

6.5 Conclusion

We have demonstrated that GEM is suitable for on-demand retrieval of independent optical channels stored in the memory. To improve the flexibility of the technique, we introduced a method allowing fast erasure (compared to the transverse decay time of the ground state coherence) of undesired subregions of an image, which in principle could be extended to a spatially (though not temporally) multiplexed quantum signal [169,170]. We observed that, under our experimental conditions, the spatially addressable resolution for readout and erasure is limited by the imaging system for the read and eraser beams and that the maximum density of information that can be spatially multiplexed is limited by the atomic diffusion of the atoms.

6.6 Attempt to Store Continuous–Variable Entanglement

While these multiplexing experiments are certainly encouraging, a more relevant question is whether the GEM technique can be provably shown to store CV bipartite entanglement. There are certainly encouraging signs that this should be possible. For instance, warm atomic ensembles have been used to store EPR states using a related spin–wave storage technique [116]. Additionally, quantum process tomography of the Λ–GEM scheme implemented in a warm vapor of ^{87}Rb demonstrated recall fidelities that exceed the no–cloning limit [171]. Motivated by these results, we attempted to store quantum images produced by four–wave mixing in ^{85}Rb and to demonstrate a recovery of entanglement (see Fig. 6.11 for a simplified representation of the experiment and Fig. 6.12 for a more detailed description of the pulsed source.).

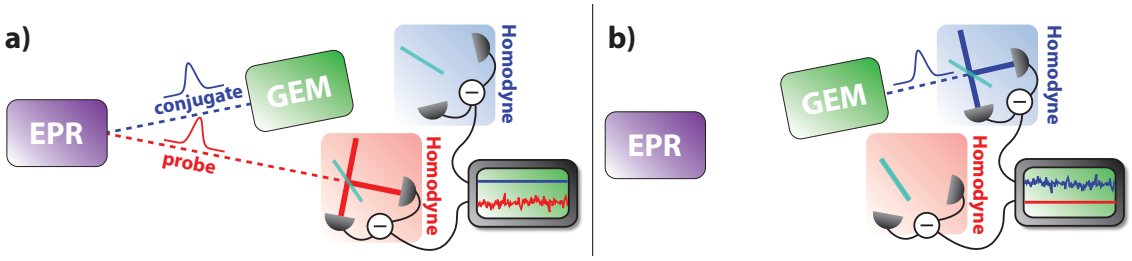


Fig. 6.11: Simple representation of an experimental attempt at storing squeezing in a Λ–GEM. a) A pulsed bipartite entangled EPR state is generated using four–wave mixing. One of the entangled modes is stored in the GEM while the fluctuations of the other mode are detected on a homodyne detector. b) After a brief storage time, the stored light is released and detected, and the measured fluctuations are compared with the previously detected fluctuations of the twin mode to check for any remaining quantum correlations.

To achieve a high degree overlap in the propagation direction of the 3 GHz detuned "control" beam and the conjugate mode being stored, we combined the two

orthogonally polarized fields on a polarizing beamsplitter. The control and conjugate fields were thus linearly and orthogonally polarized relative to one another, driving both σ^+ and σ^- transitions in the memory. The trouble is that this set–up is also nicely arranged to drive an additional four–wave mixing process in the memory cell. In the presence of phase–insensitive gain, the conjugate field being stored would be expected to suffer added noise, degrading the strength of the correlations between the twin beams.

Upon blocking the conjugate mode and leaving the control beam running continuous–wave, we observed approximately 1–2 dB of excess noise in our homodyne detection. We attributed this extra noise to the generation of a state of two–mode squeezed vacuum with the control beam on alone. This noise level was comparatively low since the memory cell was more gently heated ($\approx 80°$C) than the four–wave mixing cell ($\approx 115°$C) and since the control field was about half as intense as the four–wave mixing pump. Upon reintroducing the conjugate to the memory cell being driven by the control beam, we observed an additional 4–6 dB of added noise in the homodyne detection. Ultimately, this added noise prevented us from recovering any quantum correlations between the recalled conjugate field and its twin

One possible way around this problem might be to use same–circular polarized beams for the conjugate and control fields, which would prevent the closure of the double–Λ scheme used to achieve four–wave mixing [173]. It is presumably for this reason that the Λ–GEM experiments achieving high recall fidelities [174] combined same–circular polarizations for the stored field and the control field on a

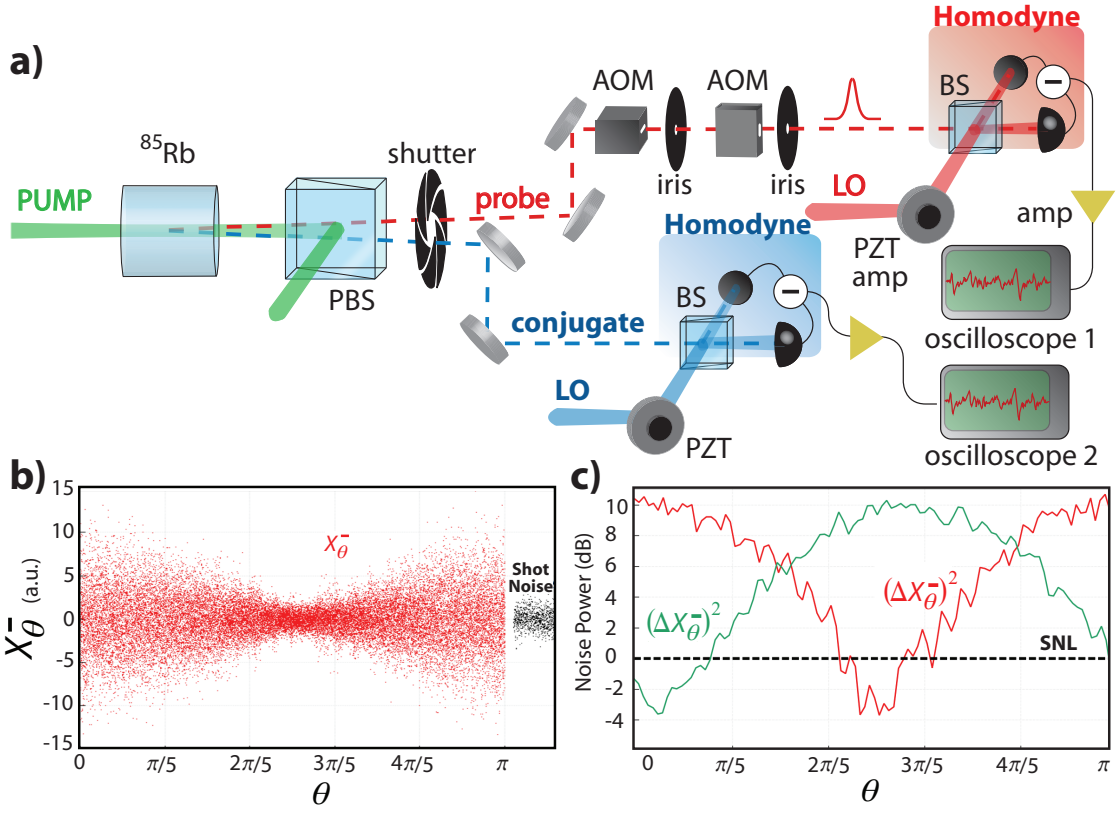


Fig. 6.12: Generation of pulsed bipartite entanglement (vacuum–squeezed twin beams) as detailed in [172]. a) The conjugate beam is sent to a homodyne detector while the probe field is pulsed using two orthogonally positioned acousto-optic modulators (AOMs). To avoid excessive loss and distortion of the mode shape, we detected the zeroth order field from the AOM. Positioning the modulators in this way achieved a greater suppression of the zeroth order. When both AOMs were simultaneously driven, we observed an extinction in optical power of a bright "test" probe field of up to 97%. b) Raw data obtained for the measurements of the joint quadrature X_θ^- as a function of the joint phase θ (detailed in [172]). c) Variance of the statistics taken in b) as a function of θ.

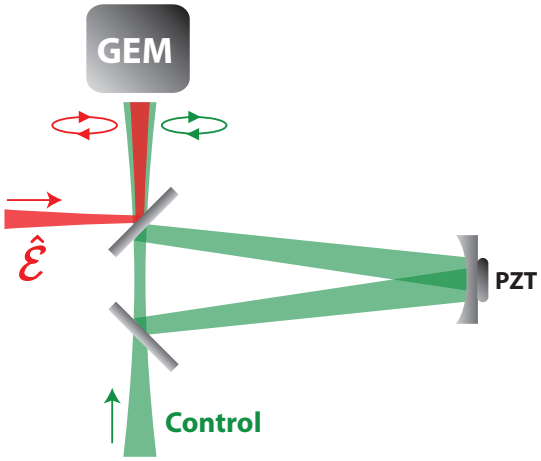


Fig. 6.13: Scheme to combine the control field and the quantum field $\hat{\mathcal{E}}$ such that they are both circularly polarized and copropagating. The cavity length is controlled using a piezoelectric transducer (PZT). The control field would need to be mode–matched to the filter cavity to achieve sufficient transmission to drive the Λ–GEM.

ring cavity. Although the cavity approach is straightforward for single–mode Raman beams[11], integrating it into our experiments would be more restrictive since we are ultimately interested in the storage of multi–spatial–mode quantum images. It should be possible to transmit the majority of the copropagating control field power through an actively stabilized resonant filter cavity while reflecting the quantum image at another color (see Fig. 6.13).

[11] Given probe and control beams in TEM_{00} modes, lenses can be used to shape the beams to be resonant with the cavity given the appropriate cavity length. In dealing with a multi–spatial quantum images, the mode overlap between the resonant cavity mode and the probe beam will be necessarily be limited.

Chapter 7: Concluding Remarks

7.1 Summary of Results

In this thesis, we studied how continuous–variable (CV) entanglement behaves when half of a bipartite entangled state is sent through a medium with anomalous dispersion associated with a phase–insensitive gain (Chapter 4). We also considered the storage and retrieval of classical images in a Gradient Echo Memory (GEM) implemented in a vapor of ^{85}Rb and studied the effect of diffusion on the images (Chapter 6). The long term objective of these experiments was to study how a rich resource of multi–spatial mode entanglement (optical four–wave mixing) can be manipulated using atomic ensembles as a dispersive medium or a storage medium.

7.1.1 Entanglement and Fast Light

Applying the formalism of quantum information theory, we studied the behavior of the entanglement and quantum mutual information of these EPR states in the presence of a gain–assisted fast–light medium. Using a powerful inseparability criterion [16] as a delay–dependent entanglement witness, we observed a statistically significant advance in the peak of the quantum correlations. Similarly, the peak of the delay–dependent quantum mutual information between the modes showed very

clear evidence of an advance. The noise added by the gain, however, caused sufficient degradation of the entanglement and mutual information that we did not observe a statistically significant advance of the leading edges of the inseparability and mutual information. In contrast, when performing similar experiments with slow light (given the same degradation of the quantum correlations), the leading and trailing edges of the inseparability and mutual information both showed significant delays.

7.1.2 Multiplexing a Gradient Echo Memory (GEM)

The multi–spatial–mode structure of the EPR states produced by 4WM might serve as a convenient source of distillable CV entanglement for a quantum repeater utilizing a multiplexed architecture. Motivated by this possibility, we considered the storage and retrieval of multiplexed optical signals using a Λ–GEM implemented in a gently heated vapor of ^{85}Rb. We showed that portions of an image written into the memory can be spatially targeted for retrieval or erasure, and we demonstrated that the memory is capable of storing multiple images simultaneously. Since these experiments were conducted in a warm atomic ensemble, we characterized the effects of atomic diffusion on the quality of the recalled images. Our conclusions were that, at temperatures near 80 °C, storage times are limited to a scale of approximately 5–10 microseconds before atomic diffusion significantly degrades the image quality. While we were unsuccessful in demonstrating any storage of CV entanglement, we found that the Λ–GEM technique provides a powerful platform to process both spatially and temporally multiplexed signals.

7.2 Outlook and Remaining Questions

7.2.1 GEM as a Quantum Memory

To me, it is still a little unclear whether a Λ–GEM in ^{85}Rb really should be expected to yield an efficient, quiet optical memory. At the time that we conducted our experiments, my intuition was that a parasitic 4WM process (discussed in the previous chapter) was primarily to blame for all observed added noise. Recently, however, a series of quantum process tomography experiments performed in ^{85}Rb at the University of Calgary produced results [173] that challenge this intuition. Using same–circular polarized light fields to drive the memory's Raman process (which would be expected to suppress any 4WM), they observed a persistence of added noise which appeared to be insensitive to any phase matching conditions. Their conclusion was that the noise was likely to result from a parasitic spontaneous Raman scattering process.

Drawing meaningful comparisons between the Calgary experiments and more optimistic conclusions [165,174–176] from the Australian National University (ANU) is difficult. For example, an important difference between the Calgary and ANU experiments was the difference in the single–photon detunings used to drive the Raman process. Specifically, the ANU experiments were conducted in ^{87}Rb using a single photon detuning of 3 GHz [175]. Our empirical observations in ^{85}Rb were consistent with the Calgary group: a 1 GHz detuning tended to yield substantially higher retrieval efficiencies. This might suggest that the ANU results benefited from

a significant decrease in spontaneous Raman scattering [134] because of their larger detuning, possibly accounting for the quieter performance of the memory.

Even with these caveats in mind, however, the ANU and Calgary results strike me as being somewhat at odds with one another. Unfortunately, our results cannot directly be brought to bear since we used orthogonal linear polarization of our Raman beams rather than matching circular polarizations. That being said, the Calgary results seem to qualitatively conform well with ours, while the ANU group reports significantly higher efficiencies. It would be interesting to understand where this apparent discrepancy emerges.

7.2.2 Technical Improvements to GEM

I was never fully satisfied with the switching performance of our time–varying magnetic field gradients used in the memory (we used commercially available devices with integrated MOSFETs). Although our 10% $\rightarrow$ 90% magnetic field switching times were comparable to those of other groups (typically on the sub–microsecond time scale), we often observed significant ringing (a modulation depth of up to 5%, though this could largely be suppressed with a flyback diode) for several microseconds after the switch, reaching into the echo recovery time[1]. I would be interested in seeing how a more diligent effort to critically damp the inductive load (or even to drive the switching with active feedback) might affect the efficiency and the distortion of the recalled echo.

Additionally, it might be worthwhile to try reversing the current in a single

[1] We measured this ringing by monitoring the voltage across a resistor placed in series with the coil generating the magnetic field.

gradient coil (as opposed to driving two oppositely wound gradient coils) to achieve the rephasing of the spin wave. Any deviation from a linear field gradient should in principle be compensated during recall if the same coil is used. One immediate disadvantage of this technique would be an average offset in the magnetic field strength, leading to a detuning of the central frequency of the recalled echo. This detuning might complicate the homodyne detection of the echo due to a mismatch between the target state and the local oscillator frequency using our typical techniques. In principle, however, this frequency mismatch could be handled by a matched frequency shift in the control beam or in the local oscillator during the detection of the echo.

7.2.3 Fast Light and Quantum Information

For me, there are interesting questions that remain regarding the behavior of CV entanglement and mutual information in fast–light media. The discussion of Fermi's problem in Chapter 2 was meant to point out how classically–rooted notions of a signal or an information velocity become a bit more nuanced with a fully quantum mechanical description of light[2]. Although I think we have shown very clear evidence that anomalous dispersion affects the transport of quantum information and entanglement, I don't think our failure to observe an advance in the leading edges proves that such an advance is impossible[3].

Additionally, it is unclear to me how these experiments fit with the picture

[2] I should gratefully acknowledge informal discussions with Dan Gauthier, who emphasized this point quite persuasively.

[3] The maximum "advance" in the averaged mutual information curves (occuring at a relative delay of $t = -101$ ns) is 0.6 ns, where the experimental uncertainty is ± 1.6 ns.

of "non–analytic points" acting as the carriers of information. Does it make sense to ask "where" the quantum information is contained in our twin beams, whose fluctuations are confined to a finite bandwidth? When does nature "decide" that the field should fluctuate up or down? Should we be allowed to observe an advance in the leading edge of the mutual information, or would this violate some deeper information principle?

Thinking along these lines, it might be worthwhile to pursue what kind of behavior would be expected when performing these experiments in a *phase–sensitive* amplifier (PSA). Since an ideal PSA acts as a single–mode squeezing operation, it should not be expected to alter the purity of the two–mode state [86]. What does that imply for the mutual information of the bipartite state? Should we expect a phase–sensitive amplifier to yield any dispersion? I think that answering these questions should be experimentally accessible with a recently demonstrated noiseless optical amplifier demonstrated using a similar four–wave mixing process in ^{85}Rb [30]. Certainly, there is plenty left to do.

Bibliography

[1] Roy J. Glauber. Coherent and incoherent states of the radiation field. *Phys. Rev.*, 131:2766, 1963.

[2] Mark Fox. *Quantum Optics: an Introduction*. Oxford University Press, New York, 2006.

[3] C. C. Gerry and Peter Knight. *Introductory Quantum Optics*. Cambridge University Press, New York, 2005.

[4] R. K. Pathria. *Statistical Mechanics*. Butterworth-Heinemann, Oxford, 1996.

[5] L. Mandel. Squeezing and photon antibunching in harmonic generation. *Opt. Commun.*, 42:437, 1982.

[6] N. J. Cerf, Gerd Leuchs, and E. S. Polzik. *Quantum Information with Continuous Variables of Atoms and Light*. Imperial College Press, London, 2007.

[7] Samuel L. Braunstein and Peter van Loock. Quantum information with continuous variables. *Rev. Mod. Phys.*, 77:513, 2005.

[8] Claude Cohen-Tannoudji, Jacques Dupont-Roc, and Gilbert Grynberg. *Atom-Photon Interactions: Basic Processes and Applications*. Wiley, New York, 1992.

[9] David J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Prentice Hall, Upper Saddle River, 2005.

[10] Ulf Leonhardt. *Measuring the Quantum State of Light*. Cambridge University Press, Cambridge, 2005.

[11] Marlan O. Scully and Muhammad Suhail Zubairy. *Quantum Optics*. Cambridge University Press, New York, 1997.

[12] Jean-Paul Blaizot and Georges Ripka. *Quantum Theory of Finite Systems*. MIT Press, Cambridge, 1986.

[13] R. Gilmore. Baker-Campbell-Hausdorff formulas. *J. Math. Phys.*, 15:2090, 1974.

[14] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Univ. Press, New York, 2010.

[15] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777, 1935.

[16] Lu-Ming Duan, G. Giedke, J. I. Cirac, and P. Zoller. Inseparability criterion for continuous variable systems. *Phys. Rev. Lett.*, 84:2722, 2000.

[17] R. E. Slusher, L. W. Hollberg, B. Yurke, J. C. Mertz, and J. F. Valley. Observation of squeezed states generated by four-wave mixing in an optical cavity. *Phys. Rev. Lett.*, 55:2409, 1985.

[18] H. Bechmann-Pasquinucci and W. Tittel. Quantum cryptography using larger alphabets. *Phys. Rev. A*, 61:062308, 2000.

[19] I.V. Sokolov, M.I. Kolobov, A. Gatti, and L.A. Lugiato. Quantum holographic teleportation. *Opt. Commun.*, 193:175, 2001.

[20] V. Boyer, A. M. Marino, and P. D. Lett. Generation of spatially broadband twin beams for quantum imaging. *Phys. Rev. Lett.*, 100:143601, 2008.

[21] E. Brambilla, A. Gatti, M. Bache, and L. A. Lugiato. Simultaneous near-field and far-field spatial quantum correlations in the high-gain regime of parametric down-conversion. *Phys. Rev. A*, 69:023802, 2004.

[22] M. Martinelli, N. Treps, S. Ducci, S. Gigan, A. Matre, and C. Fabre. Experimental study of the spatial distribution of quantum correlations in a confocal optical parametric oscillator. *Phys. Rev. A*, 67:023808, 2003.

[23] A. Gatti, E. Brambilla, M. Bache, and L. A. Lugiato. Correlated imaging, quantum and classical. *Phys. Rev. A*, 70:013802, 2004.

[24] Mikhail I. Kolobov and Claude Fabre. Quantum limits on optical resolution. *Phys. Rev. Lett.*, 85:3789, 2000.

[25] G. Brida, M. Genovese, and I. Ruo Berchera. Experimental realization of sub-shot-noise quantum imaging. *Nat. Photon*, 4:227, 2010.

[26] Jeremy B. Clark, Zhifan Zhou, Quentin Glorieux, Alberto M. Marino, and Paul D. Lett. Imaging using quantum noise properties of light. *Opt. Express*, 20:17050, 2012.

[27] A. M. Marino, J. B. Clark, Q. Glorieux, and P. D. Lett. Extracting spatial information from noise measurements of multi-spatial-mode quantum states. *Eur. Phys. J. D*, 66:1, 2012.

[28] Kirk McKenzie, Eugeniy E. Mikhailov, Keisuke Goda, Ping Koy Lam, Nicolai Grosse, Malcolm B. Gray, Nergis Mavalvala, and David E. McClelland. Quantum noise locking. *J. Opt. B: Quantum Semiclass. Opt.*, 7:421, 2005.

[29] Carlton M. Caves. Quantum limits on noise in linear amplifiers. *Phys. Rev. D*, 26:1817, 1982.

[30] N. V. Corzo, A. M. Marino, K. M. Jones, and P. D. Lett. Noiseless optical amplifier operating on hundreds of spatial modes. *Phys. Rev. Lett.*, 109:043602, 2012.

[31] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:749, 1932.

[32] Hermann Weyl. *The Theory of Groups and Quantum Mechanics.* Dover, New York, 1950.

[33] Thomas L. Curtright and Cosmas K. Zachos. Quantum mechanics in phase space. *Pac. Phys. N.*, 1:37, 2012.

[34] J. E. Moyal. Quantum mechanics as a statistical theory. *Proc. Cambridge Philos. Soc.*, 45:99, 1949.

[35] H.J. Groenewold. On the principles of elementary quantum mechanics. *Physica*, 12:405, 1946.

[36] Gerardo Adesso. *Entanglement of Gaussian States*. PhD thesis, University of Salerno, 2007.

[37] S. Olivares. Quantum optics in the phase space. *Eur. Phys. J. Spec. Top.*, 203:3, 2012.

[38] Warwick Bowen. *Experiments towards a quantum information network with squeezed light and entanglement*. PhD thesis, Australian National University, 2003.

[39] John David Jackson. *Classical Electrodynamics.* Wiley, New York, 1962.

[40] Stanley J. Farlow. *Partial Differential Equations for Scientists and Engineers*. Dover Publications, New York, 1993.

[41] T. H. Havelock. *The Propagation of Disturbances in Dispersive Media*. BiblioBazaar, Charleston, 2009.

[42] Peter W. Milonni. *Fast Light, Slow Light, and Left-handed Light*. Institute of Physics, Philadelphia, 2005.

[43] Aephraim M. Steinberg and Raymond Y. Chiao. Dispersionless, highly superluminal propagation in a medium with a gain doublet. *Phys. Rev. A*, 49:2071, 1994.

[44] H. M. Nussenzveig. *Causality and Dispersion Relations*. Academic Press, Waltham, 1972.

[45] John S. Toll. Causality and the dispersion relation: Logical foundations. *Phys. Rev.*, 104:1760, 1956.

[46] Robert W. Boyd. *Nonlinear Optics*. Academic Press, San Diego, 2008.

[47] Ram P. Kanwal. *Linear Integral Equations*. Birkhuser Boston, Berlin, 1996.

[48] David J. Griffiths. *Introduction to Electrodynamics*. Prentice Hall, Englewood Cliffs, 1989.

[49] P. Dirac. The quantum theory of the emission and absorption of radiation. *Proc. R. Soc. Lond. A*, 114:243, 1927.

[50] George M. Gehring, Aaron Schweinsberg, Christopher Barsi, Natalie Kostinski, and Robert W. Boyd. Observation of backward pulse propagation through a medium with a negative group velocity. *Science*, 312:895, 2006.

[51] L. J. Wang, A. Kuzmich, and A. Dogariu. Gain-assisted superluminal light propagation. *Nature*, 406:277, 2000.

[52] Eugene Hecht. *Optics*. Addison-Wesley, Reading, 1998.

[53] Robert W. Boyd. Slow and fast light: fundamentals and applications. *J. Mod. Opt.*, 56:1908, 2009.

[54] C. G. B. Garrett and D. E. McCumber. Propagation of a Gaussian light pulse through an anomalous dispersion medium. *Phys. Rev. A*, 1:305, 1970.

[55] Dmitry Budker. "superluminal phenomena". *Berkley.edu*, N.p. 9 Sept. 2002, Web. 11 Nov. 2013.

[56] Michael Stenner. *Measurement of the Information Velocity in Fast- and Slow-Light Optical Pulse Propagation*. PhD thesis, Duke University, 2004.

[57] W. Heitmann and G. Nimtz. On causality proofs of superluminal barrier traversal of frequency band limited wave packets. *Physics Letters A*, 196:154, 1994.

[58] Bernard F. Schutz. *A First Course in General Relativity*. Cambridge University Press, New York, 2009.

[59] Raymond Y. Chiao and Aephraim M. Steinberg. *Progress in Optics*. Elsevier, Amsterdam, 1997.

[60] L. Brillouin. Über die Fortpflanzung des Lichtes in dispergierenden Medien. *Annalen der Physik*, 349:203, 1914.

[61] A. Sommerfeld. Ein Einwand gegen die Relativtheorie der Elektrodynamik und seine Beseitigung. *Physik Z.*, 8:841, 1907.

[62] Kurt E. Oughstun and George C. Sherman. *Electromagnetic Pulse Propagation in Causal Dielectrics*. Springer, Berlin, 1997.

[63] J. M. Hickmann, M. Alencar, W. F. Silva, and S. T. Souza. Causality-induced pulse steepening and shock-like waves in superluminal media. *EPL*, 97:44009, 2012.

[64] Todd Rowland. Bump function. *Wolfram Mathworld: http://mathworld.wolfram.com/BumpFunction.html*.

[65] Michael D. Stenner, Daniel J. Gauthier, and Mark A. Neifeld. The speed of information in a fast-light optical medium. *Nature*, 425:695, 2003.

[66] Asher Peres and Daniel R. Terno. Quantum information and relativity theory. *Rev. Mod. Phys.*, 76:93, 2004.

[67] Rolf Landauer. Information is physical. *Phys. Today*, 44:23, 1991.

[68] Paul G. Kwiat, Edo Waks, Andrew G. White, Ian Appelbaum, and Philippe H. Eberhard. Ultrabright source of polarization-entangled photons. *Phys. Rev. A*, 60:773, 1999.

[69] Albert Einstein and Paul Arthur Schilpp. *Autobiographical Notes*. Open Court Pub. Co., La Salle, 1991.

[70] Yakir Aharonov, Benni Reznik, and Ady Stern. Quantum limitations on superluminal propagation. *Phys. Rev. Lett.*, 81:2190, 1998.

[71] John Maddox. Time machines still over horizon. *Nature*, 367:509, 1994.

[72] Raymond Y. Chiao, Alexander E. Kozhekin, and Gershon Kurizki. Tachyon-like excitations in inverted two-level media. *Phys. Rev. Lett.*, 77:1254, 1996.

[73] Bilha Segev, Peter W. Milonni, James F. Babb, and Raymond Y. Chiao. Quantum noise and superluminal propagation. *Phys. Rev. A*, 62:022114, 2000.

[74] Enrico Fermi. Quantum theory of radiation. *Rev. Mod. Phys.*, 4:87, 1932.

[75] Peter Milonni. Interatomic signaling in QED. *Nature*, 372:325, 1994.

[76] P. W. Milonni, D. F. V. James, and H. Fearn. Photodetection and causality in quantum optics. *Phys. Rev. A*, 52:1525, 1995.

[77] Peter Milonni, K. Furuya, and Raymond Chiao. Quantum theory of superluminal pulse propagation. *Opt. Express*, 8:59, 2001.

[78] J.D. Franson. Generation of entanglement outside of the light cone. *J. Mod. Opt.*, 55:2117, 2008.

[79] Claude Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379, 1948.

[80] Stefan M. Moser and Po-Ning Chen. *A Student's Guide to Coding and Information Theory*. Cambridge University Press, New York, 2012.

[81] T. M. Cover and J. Thomas. *Elements of Information Theory*. Wiley-Interscience, Hoboken, 2006.

[82] Ralph Hartley. Transmission of information. *Bell System Technical Journal*, 7:535, 1928.

[83] Mark Wilde. *Quantum Information Theory*. Cambridge University Press, New York, 2013.

[84] J. J. Sakurai and San Fu Tuan. *Modern Quantum Mechanics*. Addison-Wesley Pub. Co., Reading, 1994.

[85] John Von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer, New York, 1996.

[86] Alessio Serafini, Fabrizio Illuminati, and Silvio De Siena. Symplectic invariants, entropic measures and correlations of Gaussian states. *J. Phys. B: At. Mol. Opt. Phys.*, 37:21, 2004.

[87] Harold Ollivier and Wojciech H. Zurek. Quantum discord: A measure of the quantumness of correlations. *Phys. Rev. Lett.*, 88:017901, 2001.

[88] A. Ferraro, L. Aolita, D. Cavalcanti, F. M. Cucchietti, and A. Acn. Almost all quantum states have nonclassical correlations. *Phys. Rev. A*, 81:052318, 2010.

[89] Ulrich Vogl, Ryan T. Glasser, Quentin Glorieux, Jeremy B. Clark, Neil V. Corzo, and Paul D. Lett. Experimental characterization of Gaussian quantum discord generated by four-wave mixing. *Phys. Rev. A*, 87:010101, 2013.

[90] Christian Weedbrook, Stefano Pirandola, Raul Garcia-Patron, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd. Gaussian quantum information. *Rev. Mod. Phys.*, 84:621, 2012.

[91] G. S. Agarwal. Entropy, the Wigner distribution function, and the approach to equilibrium of a system of coupled harmonic oscillators. *Phys. Rev. A*, 3:828, 1971.

[92] Ryan M. Camacho, Michael V. Pack, John C. Howell, Aaron Schweinsberg, and Robert W. Boyd. Wide-bandwidth, tunable, multiple-pulse-width optical delays using slow light in cesium vapor. *Phys. Rev. Lett.*, 98:153601, 2007.

[93] Lene Vestergaard Hau, S. E. Harris, Zachary Dutton, and Cyrus H. Behroozi. Light speed reduction to 17 metres per second in an ultracold atomic gas. *Nature*, 397:594, 1999.

[94] Ryan T. Glasser, Ulrich Vogl, and Paul D. Lett. Demonstration of images with negative group velocities. *Opt. Express*, 20:13702, 2012.

[95] J. Keaveney, I. G. Hughes, A. Sargsyan, D. Sarkisyan, and C. S. Adams. Maximal refraction and superluminal propagation in a gaseous nanolayer. *Phys. Rev. Lett.*, 109:233001, 2012.

[96] P. W. Milonni. *Fast Light, Slow Light and Left-Handed Light*. CRC Press, Philadelphia, 2010.

[97] A. Kuzmich, A. Dogariu, L. J. Wang, P. W. Milonni, and R. Y. Chiao. Signal velocity, causality, and quantum noise in superluminal light pulse propagation. *Phys. Rev. Lett.*, 86:3925, 2001.

[98] Vincent Boyer, Alberto M. Marino, Raphael C. Pooser, and Paul D. Lett. Entangled images from four-wave mixing. *Science*, 321:544, 2008.

[99] Asher Peres and Daniel R. Terno. Quantum information and relativity theory. *Rev. Mod. Phys.*, 76:93, 2004.

[100] Samuel L. Braunstein and Peter van Loock. Quantum information with continuous variables. *Rev. Mod. Phys.*, 77:513, 2005.

[101] Alexander I. Lvovsky, Barry C. Sanders, and Wolfgang Tittel. Optical quantum memory. *Nat. Photon*, 3:706, 2009.

[102] A. M. Marino, R. C. Pooser, V. Boyer, and P. D. Lett. Tunable delay of EinsteinPodolskyRosen entanglement. *Nature*, 457:859, 2009.

[103] Robert W. Boyd. Slow and fast light: fundamentals and applications. *J. Mod. Opt.*, 56:1908, 2009.

[104] D.C. Hutchings, M. Sheik-Bahae, D.J. Hagan, and E.W. Stryland. Kramers-Kronig relations in nonlinear optics. *Opt. Quantum Electron.*, 24:1, 1992.

[105] Ulrich Vogl, Ryan T. Glasser, Jeremy B. Clark, Quentin Glorieux, Tian Li, Neil V. Corzo, and Paul D. Lett. Advanced quantum noise. *arXiv:1305.6614*, 2013.

[106] Robert W. Boyd, Zhimin Shi, and Peter W. Milonni. Noise properties of propagation through slow- and fast-light media. *J. Opt.*, 12:104007, 2010.

[107] Ryan T. Glasser, Ulrich Vogl, and Paul D. Lett. Stimulated generation of superluminal light pulses via four-wave mixing. *Phys. Rev. Lett.*, 108:173902, 2012.

[108] Michael D. Stenner, Daniel J. Gauthier, and Mark A. Neifeld. Fast causal information transmission in a medium with a slow group velocity. *Phys. Rev. Lett.*, 94:053902, 2005.

[109] Michael C. Parker and Stuart D. Walker. Information transfer and Landauers principle. *Opt. Commun.*, 229:23, 2004.

[110] Robert W. Boyd and Paul Narum. Slow- and fast-light: fundamental limitations. *J. Mod. Opt.*, 54:2403, 2007.

[111] M. D. Lukin. Colloquium: Trapping and manipulating photon states in atomic ensembles. *Rev. Mod. Phys.*, 75:457, 2003.

[112] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller. Long-distance quantum communication with atomic ensembles and linear optics. *Nature*, 414:413, 2001.

[113] Mikael Afzelius, Christoph Simon, Hugues de Riedmatten, and Nicolas Gisin. Multimode quantum memory based on atomic frequency combs. *Phys. Rev. A*, 79:052329, 2009.

[114] Brian Julsgaard, Jacob Sherson, J. Ignacio Cirac, Jaromr Fiurek, and Eugene S. Polzik. Experimental demonstration of quantum memory for light. *Nature*, 432:482, 2004.

[115] Georg Heinze, Christian Hubrich, and Thomas Halfmann. Stopped light and image storage by electromagnetically induced transparency up to the regime of one minute. *Phys. Rev. Lett.*, 111:033601, 2013.

[116] K. Jensen, W. Wasilewski, H. Krauter, T. Fernholz, B. M. Nielsen, M. Owari, M. B. Plenio, A. Serafini, M. M. Wolf, and E. S. Polzik. Quantum memory for entangled continuous-variable states. *Nat. Phys*, 7:13, 2011.

[117] G. Hetet, J. J. Longdell, A. L. Alexander, P. K. Lam, and M. J. Sellars. Electro-optic quantum memory for light using two-level atoms. *Phys. Rev. Lett.*, 100:023601, 2008.

[118] Alexey V. Gorshkov, Axel Andre, Mikhail D. Lukin, and Anders S. Srensen. Photon storage in -type optically dense atomic media. i. cavity model. *Phys. Rev. A*, 76:033804, 2007.

[119] Alexey V. Gorshkov, Axel Andre, Mikhail D. Lukin, and Anders S. Srensen. Photon storage in -type optically dense atomic media. II. free-space model. *Phys. Rev. A*, 76:033805, 2007.

[120] Alexey V. Gorshkov, Axel Andre, Mikhail D. Lukin, and Anders S. Sorensen. Photon storage in -type optically dense atomic media. III. effects of inhomogeneous broadening. *Phys. Rev. A*, 76:033806, 2007.

[121] Irina Novikova, Alexey V. Gorshkov, David F. Phillips, Anders S. Srensen, Mikhail D. Lukin, and Ronald L. Walsworth. Optimal control of light pulse storage and retrieval. *Phys. Rev. Lett.*, 98:243602, 2007.

[122] Alexey V. Gorshkov, Axel Andre, Michael Fleischhauer, Anders S. Srensen, and Mikhail D. Lukin. Universal approach to optimal photon storage in atomic media. *Phys. Rev. Lett.*, 98:123601, 2007.

[123] M. Fleischhauer and M. D. Lukin. Dark-state polaritons in electromagnetically induced transparency. *Phys. Rev. Lett.*, 84:5094, 2000.

[124] M. Fleischhauer and M. D. Lukin. Quantum memory for photons: Dark-state polaritons. *Phys. Rev. A*, 65:022314, 2002.

[125] Mahdi Hosseini, Ben M. Sparkes, Gabriel Hetet, Jevon J. Longdell, Ping Koy Lam, and Ben C. Buchler. Coherent optical pulse sequencer for quantum applications. *Nature*, 461:241, 2009.

[126] M. Hosseini, B. M. Sparkes, G. T. Campbell, P. K. Lam, and B. C. Buchler. Storage and manipulation of light using a raman gradient-echo process. *J. Phys. B: At. Mol. Opt. Phys.*, 45:124004, 2012.

[127] G. Campbell, M. Hosseini, B. M. Sparkes, P. K. Lam, and B. C. Buchler. Time- and frequency-domain polariton interference. *New J. Phys.*, 14:033022, 2012.

[128] Peter Lambropoulos and David Petrosyan. *Quantum Optics and Quantum Information*. Springer, New York, 2007.

[129] M. Gross and S. Haroche. Superradiance: An essay on the theory of collective spontaneous emission. *Physics Reports*, 93:301, 1982.

[130] R. H. Dicke. Coherence in spontaneous radiation processes. *Phys. Rev.*, 93:99, 1954.

[131] Marlan O. Scully and Anatoly A. Svidzinsky. The super of superradiance. *Science*, 325:1510, 2009.

[132] Michael Fleischhauer and M.D. Lukin. Quantum Memory for Photons: Dark-state Polaritons. *Phys. Rev. A*, 65:22314, 2002.

[133] S. A. Moiseev and S. Kroll. Complete reconstruction of the quantum state of a single-photon wave packet absorbed by a doppler-broadened transition. *Phys. Rev. Lett.*, 87:173601, 2001.

[134] C. J. Foot. *Atomic Physics*. Oxford University Press, Oxford, 2005.

[135] B. Kraus, W. Tittel, N. Gisin, M. Nilsson, S. Kroll, and J. I. Cirac. Quantum memory for nonstationary light fields based on controlled reversible inhomogeneous broadening. *Phys. Rev. A*, 73:020302, 2006.

[136] Daniel A. Steck. "Rubidium 85 D Line Data". *available online at https://steck.us/alkalidata*, N.p. 20 Sept. 2013, Web. 19 Nov. 2013.

[137] Michael Fleischhauer, Atac Imamoglu, and Jonathan P. Marangos. Electromagnetically induced transparency: Optics in coherent media. *Rev. Mod. Phys.*, 77:633, 2005.

[138] Mahdi Hosseini. *Quantum Optical Storage and Processing Using Raman Gradient Echo Memory*. PhD thesis, Australian National University, 2012.

[139] H. J. Kimble. The Quantum Internet. *Nature*, 453:1023, 2008.

[140] Fabrizio Illuminati. Quantum optics: Light does matter. *Nat. Phys.*, 2:803, 2006.

[141] Antonio Acn, J. Ignacio Cirac, and Maciej Lewenstein. Entanglement percolation in quantum networks. *Nat. Phys.*, 3:256, 2007.

[142] David d'Enterria and Gustavo G. da Silveira. Observing light-by-light scattering at the large hadron collider. *Phys. Rev. Lett.*, 111:080405, 2013.

[143] H.-J. Briegel, W. Dr, J. I. Cirac, and P. Zoller. Quantum repeaters: The role of imperfect local operations in quantum communication. *Phys. Rev. Lett.*, 81:5932, 1998.

[144] John Preskill. Fault-tolerant quantum computation. *arXiv quant-ph/9712048*, 1997.

[145] David Deutsch, Artur Ekert, Richard Jozsa, Chiara Macchiavello, Sandu Popescu, and Anna Sanpera. Quantum privacy amplification and the security of quantum cryptography over noisy channels. *Phys. Rev. Lett.*, 77:2818, 1996.

[146] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Phys. Rev. Lett.*, 76:722, 1996.

[147] Thomas Jennewein, Gregor Weihs, Jian-Wei Pan, and Anton Zeilinger. Experimental nonlocality proof of quantum teleportation and entanglement swapping. *Phys. Rev. Lett.*, 88:017903, 2001.

[148] Jian-Wei Pan, Matthew Daniell, Sara Gasparoni, Gregor Weihs, and Anton Zeilinger. Experimental demonstration of four-photon entanglement and high-fidelity teleportation. *Phys. Rev. Lett.*, 86:4435, 2001.

[149] Charles H. Bennett, Gilles Brassard, Claude Crepeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895, 1993.

[150] M. Zukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert. Event-ready-detectors bell experiment via entanglement swapping. *Phys. Rev. Lett.*, 71:4287, 1993.

[151] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824, 1996.

[152] Christoph Simon, Hugues de Riedmatten, Mikael Afzelius, Nicolas Sangouard, Hugo Zbinden, and Nicolas Gisin. Quantum repeaters with photon pair sources and multimode memories. *Phys. Rev. Lett.*, 98:190503, 2007.

[153] M. Bonarota, J.-L. Le Gouet, and T. Chaneliere. Highly multimode storage in a crystal. *New J. Phys.*, 13:013013, 2011.

[154] O. Collins, S. Jenkins, A. Kuzmich, and T. Kennedy. Multiplexed Memory-Insensitive Quantum Repeaters. *Phys. Rev. Lett.*, 98:060502, 2007.

[155] S. Y Lan, A. G. Radnaev, O. A. Collins, D. N. Matsukevich, T. A. B. Kennedy, and A. Kuzmich. A multiplexed quantum memory. *Opt. Express*, 17:13639, 2009.

[156] M. Shuker, O. Firstenberg, R. Pugatch, A. Ron, and N. Davidson. Storing images in warm atomic vapor. *Phys. Rev. Lett.*, 100:223601, 2008.

[157] O. Firstenberg, M. Shuker, R. Pugatch, D. R. Fredkin, N. Davidson, and A. Ron. Theory of thermal motion in electromagnetically induced transparency: Effects of diffusion, Doppler broadening, and Dicke and Ramsey narrowing. *Phys. Rev. A*, 77:043830, 2008.

[158] F. A. Franz. Rubidium spin relaxation in the rare gases under ultraclean conditions. *Phys. Rev.*, 139:603, 1965.

[159] W. Franzen. Spin relaxation of optically aligned rubidium vapor. *Phys. Rev.*, 115:850, 1959.

[160] L. Zhao, T. Wang, Y. Xiao, and S. F. Yelin. Image storage in hot vapors. *Phys. Rev. A*, 77:041802, 2008.

[161] Quentin Glorieux, Jeremy B. Clark, Alberto M. Marino, Zhifan Zhou, and Paul D. Lett. Temporally multiplexed storage of images in a gradient echo memory. *Opt. Express*, 20:12350, 2012.

[162] Praveen K. Vudyasetu, Ryan M. Camacho, and John C. Howell. Storage and retrieval of multimode transverse images in hot atomic rubidium vapor. *Phys. Rev. Lett.*, 100:123903, 2008.

[163] M. Shuker, O. Firstenberg, R. Pugatch, A. Ron, and N. Davidson. Storing Images in Warm Atomic Vapor. *Phys. Rev. Lett.*, 100:223601, 2008.

[164] D. Higginbottom, B. Sparkes, M. Rancic, O. Pinel, M. Hosseini, P. Lam, and B. Buchler. Spatial-mode storage in a gradient-echo memory. *Phys. Rev. A*, 86:023801, 2012.

[165] Mahdi Hosseini, Ben M Sparkes, Gabriel Hetet, Jevon J. Longdell, Ping Koy Lam, and Ben C. Buchler. Coherent optical pulse sequencer for quantum applications. *Nature*, 461:241, 2009.

[166] O. Firstenberg, M. Shuker, R. Pugatch, D.R. Fredkin, N. Davidson, and A. Ron. Theory of thermal motion in electromagnetically induced transparency: Effects of diffusion, Doppler broadening, and Dicke and Ramsey narrowing. *Phys. Rev. A*, 77:043830, 2008.

[167] O. Firstenberg, M. Shuker, A. Ron, and N. Davidson. Colloquium: Coherent diffusion of polaritons in atomic media. *Rev. Mod. Phys.*, 85:941, 2013.

[168] M. Hosseini, B. M. Sparkes, G. T. Campbell, P. K. Lam, and B. C. Buchler. Storage and Manipulation of Light using a Raman Gradient-echo Process. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 45:124004, 2012.

[169] Jeremy B. Clark, Zhifan Zhou, Quentin Glorieux, Alberto M. Marino, and Paul D. Lett. Imaging using quantum noise properties of light. *Opt. Express*, 20:17050, 2012.

[170] Vincent Boyer, Alberto M. Marino, Raphael C. Pooser, and Paul D. Lett. Entangled Images from Four-Wave Mixing. *Science*, 321:544, 2008.

[171] M. Hosseini, G. Campbell, B. M. Sparkes, P. K. Lam, and B. C. Buchler. Unconditional room-temperature quantum memory. *Nat. Phys.*, 7:794, 2011.

[172] Quentin Glorieux, Jeremy B. Clark, Neil V. Corzo, and Paul D. Lett. Generation of pulsed bipartite entanglement using four-wave mixing. *New J. Phys.*, 14:123024, 2012.

[173] Connor Kupchak. *Complete Characterization of Quantum Optical Porcesses with a Focus on Quantum Memory*. PhD thesis, University of Calgary, 2013.

[174] M. Hosseini, G. Campbell, B. M. Sparkes, P. K. Lam, and B. C. Buchler. Unconditional Room-temperature Quantum Memory. *Nature Physics*, 7:794, 2011.

[175] M. Hosseini, B. M. Sparkes, G. Campbell, P. K. Lam, and B. C. Buchler. High efficiency coherent optical memory with warm rubidium vapour. *Nature Communications*, 2:174, 2011.

[176] B. M. Sparkes, J. Bernu, M. Hosseini, J. Geng, Q. Glorieux, P. A. Altin, P. K. Lam, N. P. Robins, and B. C. Buchler. Gradient echo memory in an ultra-high optical depth cold atomic ensemble. *New J. Phys.*, 15:085027, 2013.